PhD project funded by:
Fundação para a ciência e tecnologia, Ministério da Ciência, tecnologia e ensino superior – Portugal (ref: SFRH/BD/60489/2009)

Characterization of the cis-regulatory regions governing *ojoplano locus*

Inês Gago Rodrigues
Director: Juan Ramón Martinez Morales
Centro Andaluz de Biologia del Desarrollo
Universidad Pablo de Olavide/ CSIC

October 2013

Contents

Tables and figures

I. Introduction

1. Regulation of Developmental Genes

During the embryogenesis of multicellular organisms, developmental genes control the regulatory networks responsible for directing the morphogenesis of the different tissues. The transcriptional regulation of these genes requires a precise orchestration of *non*-coding cis-regulatory elements (NCREs), which are genomic sequences often conserved throughout the entire vertebrate's lineage.

NCREs can be active or inactive elements, depending on their interactions with proteins and promoter sequences present in its genomic environment. These elements are found in the same DNA molecule of the gene that they regulate and contain specific binding sites for transcription factors that regulate gene activation or inhibition (Gomez-Skarmeta et al., 2006).

Comparative genomic analyses have revealed recently a collection of approximately 3000 NCREs in vertebrates (Bejerano et al., 2004; Sandelin et al., 2004); the majority of these regions have been found to be in the vicinity of important morphogenetic genes, with essential functions during development (Sandelin et al., 2004; Woolfe et al., 2005). Additionally, functional studies have reported that a large proportion of the NCREs contain either enhancer o silencer elements that regulate the expression of developmental genes in specific domains of the embryo (Gomez Skarmeta, 2009). These findings indicate that the regulatory elements hidden in the non-coding genome play an essential role in gene expression regulation during development.

In this work, I have focussed in the identification of NCREs found in the landscape of a developmental gene called *ojoplano* and in the potential of these NCREs to act as enhancers, activating gene expression.

2. The role of enhancers during embryonic development

Enhancers are major determinants of differential spatio-temporal expression. In other words, they determine where and when a gene promoter will be used, and how much of the gene product will be synthesized. Each of these regulatory elements interact physically with a specific gene promoter to control the efficiency and rate of transcription; they can be specific for one gene or for a group of genes. Additionally, they can activate transcription independently of their location, distance or orientation with respect to the promoter (Banerji et al., 1981); Figure 1).

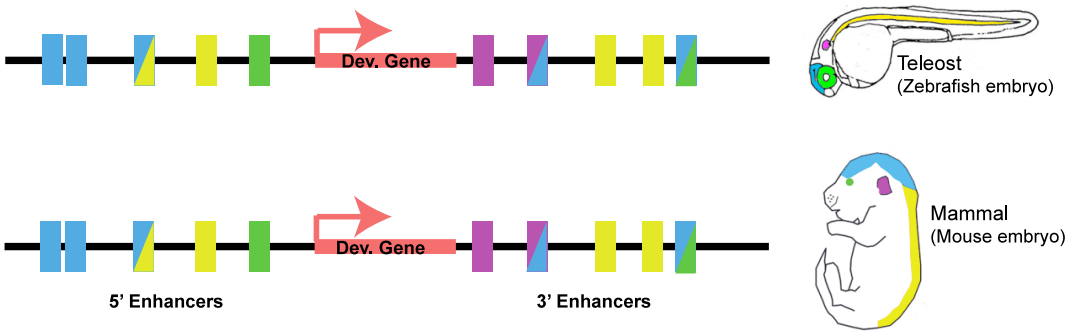


Figure 1. Tissue-specific gene expression is regulated by enhancer activity. Conserved enhancers are represented as rectangular boxes along the DNA sequence (black line) of tow different species: zebrafish (*Danio rerio*) and mouse (*Mus musculus*). The enhancers are distributed upstream and downstream from the gene they regulate (red arrow). Each color represents the activity of an enhancer in a specific domain of the embryo. Enhancers activate the transcription in a combinatorial way (multiple enhancers in the same domains of the embryo) and its activity can be modulated in one or other direction (two-color enhancers).

These regulatory regions are enriched in transcription factor binding sites, which bind to enhancer or promoter regions, activating or repressing the synthesis of mRNA from a particular gene. Transcription factors (TFs) include three major protein domains: the DNA-binding domain that recognizes a specific sequence in the DNA, the trans-activating domain that activates or represses transcription and finally a protein-protein interaction domain that allows the transcription factor activity to be modulated by other TFs and proteins. TFs within the same family share a common framework in their DNA binding sites (Graf and Enver, 2009).

Interestingly, one TF can bind to different enhancers activating the expression of different genes in different cell types. In this sense, it is the combination of several transcription factors acting on several enhancers that causes particular genes to be transcribed. Moreover, a gene can have numerous enhancer elements, each of them turning the gene on in a specific subset of cells (Graf and Enver, 2009); Figure 2).

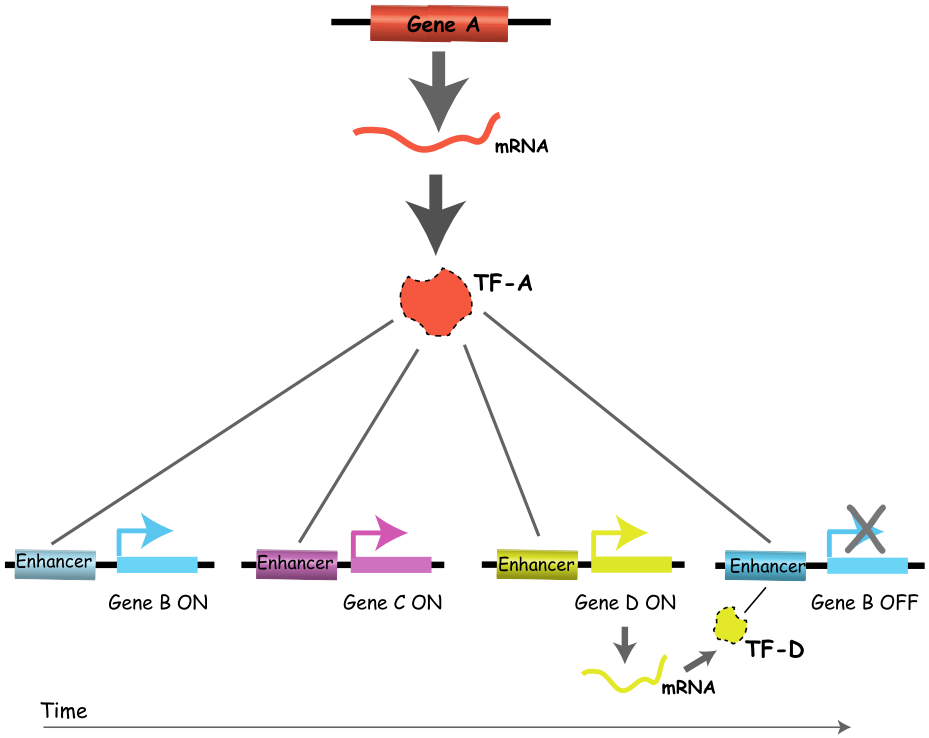


Figure 2. Transcription factors regulate gene expression through enhancer activation/inactivation. Gene A transcribes a transcription factor (TF-A). TF-A turns on genes B, C and D in a particular order. Gene D also encodes a transcription factor, TF-D. Later on in development, TF-A together with TF-D is necessary to stop the expression of gene B.

During gene expression activation, the promoter must be accessible to enhancers, TFs and other protein complexes, and this requires a decompaction of the chromatin fiber. Chromatin is constituted by nucleosomes, which are stretches of DNA, 147 base-pair (bp) long, wrapped around proteins called core histones. In the nuclesomes, core histones form an octamer, which consists of two copies of each core histone (H3, H4, H2A and H2B). The N-terminal tails of core histones are highly conserved between species and can be subjected to post-translational modifications, such as methylation, acetylation and ubiquitylation, which are additions of methyl, acetyl or ubiquitin groups onto the different lysines of histone H3 tail (Bannister and Kouzarides, 2011).
For the regulation of gene expression is required the generation of a nuclesome-free region nearby the gene promoter, which is associated with the distinct modifications in the histone tails. In fact, the localization of histone modification marks usually indicates the position of enhancers and promoters. Active enhancers can be

predicted by the monomethylation of histone H3 in lysine 4 (H3K4me1) and by the acetylation of the histone H3 at lysine 27 (H3K27ac), whereas inactive enhancers only show H3K4me1 mark. The promoters are marked with the trimethylation of histone H3 on lysine 4 (H3K4me3) and normally this histone modification marks transcription activation and subsequent gene expression (Heintzman et al., 2007; Heintzman et al., 2009; Creyghton et al., 2010); Figure 3).

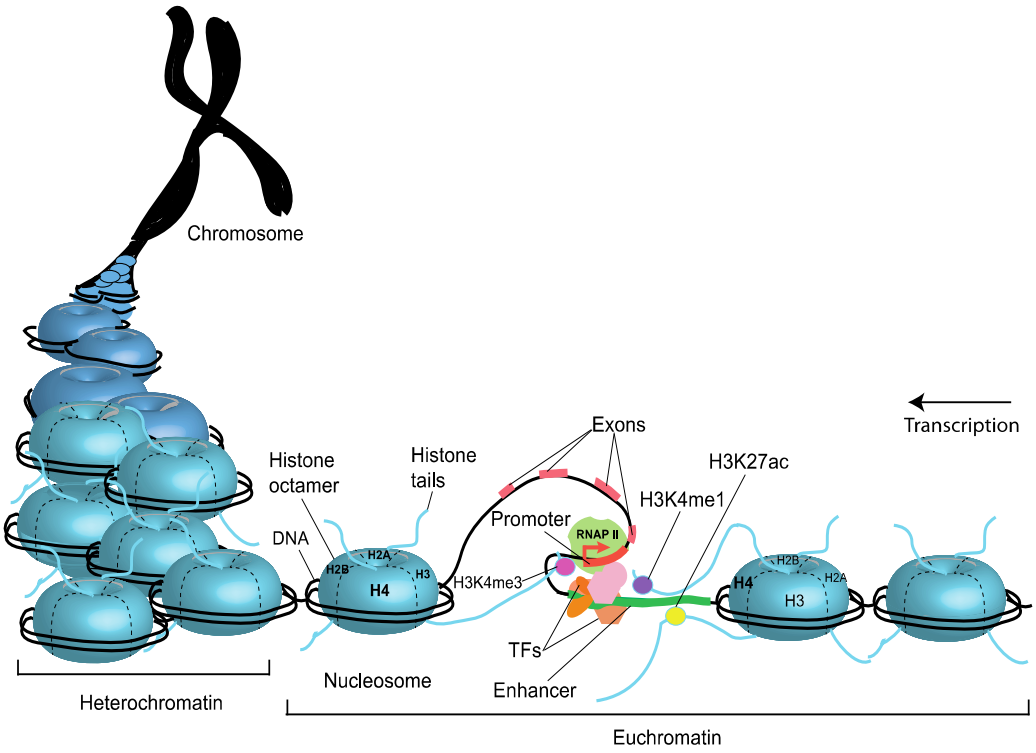


Figure 3. Transcription initiation process. The interactions between histone modification marks, enhancer-TFs and the RNA polymerase II transcription initiation protein complex (RNAP II) are required during gene transcription. For gene expression activation, chromatin must change its conformation from a compact fiber (heterochromatin) to an open fiber (euchromatin), which allows proteins and enhancers to access the gene promoter (red). During gene expression initiation, enhancers (green) marked with H3K4me1 (purple circle) and H3K27ac (yellow circle) are in contact with specific TFs (orange and light pink forms) and together they bind to the gene promoter (red), marked with H3K4me3 (dark pink circle). These interactions favor the binding of the RNAP II protein complex (light green) to the promoter and the transcription is then initiated.

New enhancers can be identified in the non-coding genome through enhancer-screenings (Bessa and Gomez-Skarmeta, 2011). These screenings scan the genome on

the bases of sequence conservation and predictive epigenetic marks, elucidating not only the regulation of specific developmental genes, but also uncovering new developmental functions associated to hidden elements in the genome.

Enhancer-screenings can also contribute to the localization of regulatory regions controlling the activity of genes associated with human hereditary diseases, which are frequently developmental genes. These type of diseases caused by deletions, duplications, translocations or point mutations, can occur in either the coding or non-coding genome, as is the case for enhancers.

In this work we study the developmental gene *Ojoplano (opo)*, which plays a crucial role during eye morphogenesis in fish (Martinez-Morales et al., 2009) and has been associated with human craniofacial hereditary diseases and schizophrenia (Box 1). We have performed an enhancer screening in the *opo locus* to investigate its regulation and functions during eye morphogenesis and to find out if any of its enhancers is localized in proximity to genetic markers previously associated to hereditary diseases (orofacial clefting and schizophrenia).

However, before discussing the regulation of *opo* during early eye morphogenesis, I will describe below the basic principles behind this developmental event.

Box 1: *opo* and human disease

Opo is a highly conserved developmental gene present in a single copy in vertebrate genomes. Its human ortholog (*ofcc1*) is located at the distal region of chromosome 6p (6p23.4), in a syntenic gene desert of approximately 2 Megabases (Mb). This locus represents only 2% of the chromosome, however it contains 23% of all its *non*-coding *cis*-regulatory elements, which may suggest the existence of complex regulatory logic in the region.

The *opo locus* in humans is associated with hereditary congenital diseases. Breakpoint mutations in patients with orofacial cleft syndrome have been associated with *ofcc1* (Davies et al., 2004). This congenital craniofacial defect is characterized by a deficient development of the palate that leads to the formation of clefts in both the lip and palate (Davies et al., 2004). Additionally, it has been shown that several genetic markers localized in the *ofcc1* region are associated to schizophrenia (Straub et al., 2002). Finally, patients with Tourette syndrome, a common neuro-developmental disorder, revealed the presence of a novel variant in the 5' untranslated region of the *ofcc1* gene (Sundaram et al., 2011).

3. Early eye development: patterning and morphogenesis of the optic cup

The developing brain can be subdivided into three anatomically distinct areas: the forebrain, the midbrain and the hindbrain. The eyes originate from a specialized region of the forebrain called the eye field (Adelmann, 1929).

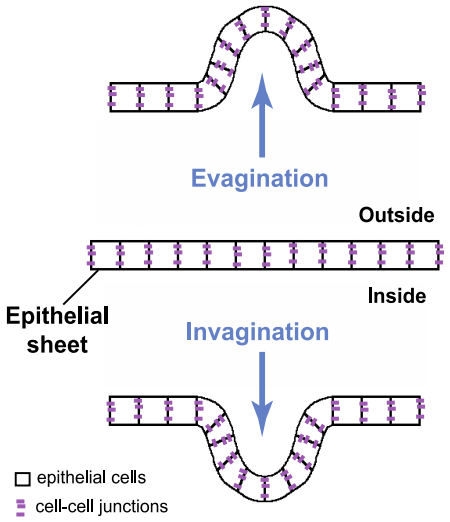


Figure 4. Evagination and invagination processes of an epithelial sheet of cells during development (adapted from http://www.aubrun.edu).

The eye field is composed of neuroepithelial cells that adhere each other through cell-cell junctions forming sheets (Figure 4). These cells have apical-basal polarity defined by distinct 'apical', 'lateral' and 'basal' plasma membrane domains. The apical membrane of the epithelial cells faces the future ventricular surface, the lateral membrane connect the neighboring cells and the basal membrane attach the cells to the basement membrane, which is a thin sheet of extracellular matrix proteins that separates the epithelial sheet from underlying cells (Bryant and Mostov, 2008). Normally epithelial sheets form cavities and surfaces during development along the animal body by evagination and invagination mechanisms (Figure 4).

The neuroepithelial cells from the eye field interact with the surrounding surface ectoderm and the extraocular mesenchyme, which have, respectively, originated from the neural crest and the mesoderm. Following the appearance of the eye field, the cells evaginate from both sides of the forebrain, resulting in the formation of the both bilateral optic vesicles (Figure 5-A).

Each optic vesicle then invaginates at its distal portion to form the optic cup (see figure 5-B and C), with the presumptive neural retina (NR) and retinal pigmented epithelium (RPE) forming the inner and outer walls. The most proximal part develops into the optic stalk, which is the structure that connects the optic vesicle to the

forebrain (Figure 5-D). At the same time, the surface ectoderm adjacent to the retina also invaginates and develops into the lens vesicle, while the rest of the surface ectoderm near the lens becomes the corneal epithelium. Following these morphogenetic events, the RPE and NR cells become specified and also spatially organized (Hilfer, 1983; Chuang and Raymond, 2002; Fuhrmann, 2010).

In summary, the early eye development can be organized in 4 steps: eye field establishment, optic vesicle evagination, optic vesicle patterning (RPE optic stalk and NR) and optic cup morphogenesis. I will next describe these steps in further detail.

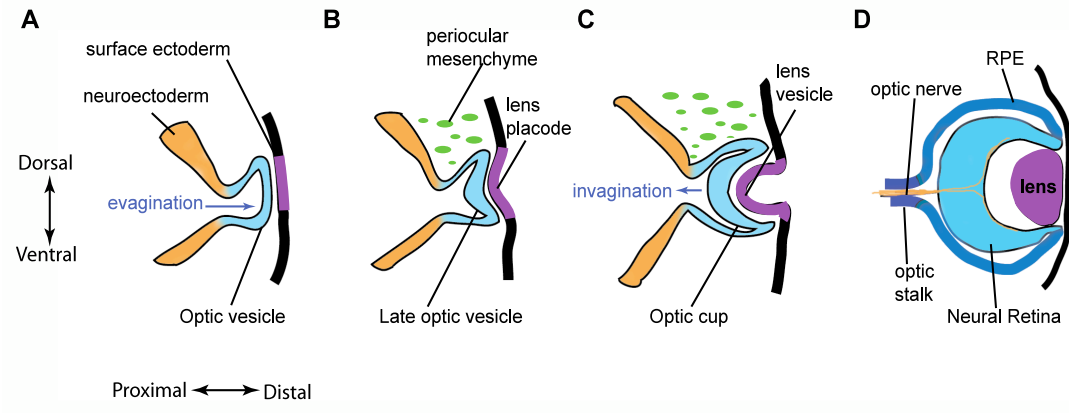


Figure 5. Representation of early eye development in mouse embryo, from E9 to E10.5. A) The optic vesicles (turquoise) evaginate from the forebrain neuroectoderm (orange) and contact with the surface ectoderm (purple). B) and C) the optic vesicles contact with the lens placode and both tissues invaginate to form the optic cup and the lens vesicles. D) After the folding of the optic cup the eye morphogenesis is completed. The optic cup is formed by the presumptive NR (light blue) at its distal part, the RPE (blue), which covers the NR, at its proximal part; the optic stalk (dark blue), which connects the optic vesicle to the forebrain through the optic nerve (orange) is localized at the most proximal part (Adapted from Eiraku et al., 2011).

Step 1: Eye field specification

After the extension of the embryonic body axis, by the convergence-extension movements, the neural plate appears as a regionalized tissue, whit the eye field located in the most anterior part (Wilson and Houart, 2004). The eye field specification and subsequent development into two bilateral optic vesicles depends on the expression of highly conserved homeobox transcription factors; a large group

of genes that share a common DNA sequence (homedomain) (Adler and Canto-Soler, 2007). The homeobox transcription factors: Otx2, Six3, Rx, Pax6, Six6 and Tbx3 are expressed in a dynamic, overlapping pattern in the presumptive eye field (Graw, 2010).

Also expressed in the eye field Sonic hedgehog (Shh) is a ligand of the hedgehog signaling pathway, which plays a key role in the regulation of the tissue growth during organogenesis. The Shh ligand and the TF Six3 are main players during the splitting of the eye field in two optic vesicles (Geng et al., 2008). Geng et al. (2008) showed that Six3 regulates Shh. The loss of either Six3 or Shh leads to holoprosencephaly with cyclopia in zebrafish embryos. Thus showing that without these genes expression the splitting of the eye field into two optic vesicles gets compromised (Geng et al., 2008).

However, eye filed splitting is a complex process that depends also on the interaction between several signaling pathways and tissues. The eye field is surrounded by telencephalic precursors and by cells that will form the hypothalamus (Esteve and Bovolenta, 2006). The prechordal plate mesoderm, which is a cell aggregate located in the anterior end of the notochord, plays a key role in determining the eye field and its subdivision into two separate domains, the future bilateral optic vesicles (Varga et al., 1999). Previous studies reveled that the removal of the prechordal plate mesoderm, which is specified by *One eye pinhead* gene (Schier et al., 1997), led to the formation of a single retina both in chick embryos and Xenopus explants (Li et al., 1997).

Step 2: Optic vesicle evagination

Optic vesicle emergence is a conserved process among vertebrates with many common features amongst species. Teleosts are, however, an exception; as in this group the formation of the optic vesicles is initiated in a different way. Instead of evaginating, as described before in the mouse, fish retinal progenitor cells individually migrate out laterally from the neural tube and form a vesicle by local epithelialization (Rembold et al., 2006).

In both, mammals and teleosts, the optic vesicle formation requires coordinated changes in cell shape and in cellular behavior, the retinal homeodomain transcription factor Rx/RAX is known to modulated these cell behavior (Rembold et al., 2006).

Previous work in medaka and zebrafish revealed that *rx3* is essential optic vesicle evagination (Winkler et al., 2000; Loosli et al., 2003; Stigloher et al., 2006); Figure 5-A). Additionally, Rx/RAX controls neuroepithelial cells movements by activating the non-canonical Wnt signaling pathway, which have been involved in the specification of the eye field in vertebrates (Cavodeassi et al., 2005; Martinez-Morales and Wittbrodt, 2009). Finally, there is evidence also from zebrafish, showing that Rx regulates cell proliferation during optic vesicle evagination (Stigloher et al., 2006; Fuhrmann, 2010). These results indicate the Rx expression is essential for eye field specification and the subsequent evagination of the optic vesicles.

After evagination, neuroepithelial cells are still morphologically and molecularly indistinguishable but do have the capability to develop into different cell types. During optic vesicle patterning multiple signaling molecules act to modify the identity of the cells, and define their differentiation into NR, optic stalk or RPE cells (Martinez-Morales et al., 2004). I will summarize the molecular interactions responsible for the patterning of the optic vesicles in subdomains below.

Step 3: Optic vesicle patterning

The undifferentiated cells from the optic vesicles express a conserved group of TFs, including Rx, Pax6, Hes1, Otx2, Lhx2, Six3, and Six9 (Adler and Canto-Soler, 2007). These genes, which form the nodes of the eye-specific regulatory network, are also responsible for the segregation of the different subdomains (NR, optic stalk and RPE) in the vesicle. (Figure 6).

The main molecular network involved in the patterning of the optic vesicle starts with the LIM homeobox transcription factor Lhx2, which is the earliest patterning gene known for the optic vesicle (Fuhrmann, 2010). This gene is responsible for the specification of both the RPE and the NR (Yun et al., 2009). It activates the bHLH transcription factor Mitf in presumptive RPE and the homeobox transcription factor Vsx2 in the presumptive NR (Burmeister et al., 1996; Green et al., 2003). Initially, *mitf* is expressed throughout the optic vesicle. But as development proceeds, it is

subsequently downregulated at the distal domain of the optic vesicles, where *vsx2* expression is initiated (Nguyen and Arnheiter, 2000). The proximal domain of the optic vesicle, which connects the optic vesicles to forebrain, expresses *pax2* and is thus specified into optic stalk. Previous experiments with *pax2*$^{-/-}$ mice revealed that the interaction between Pax2 and Pax6 is responsible for the establishment of the boundaries between the optic vesicles and the optic stalk (Schwarz et al., 2000) (Figure 6).

During the optic vesicle specification the prospective NR and the RPE are in contact with the surface ectoderm and the periocular mesenchyme, respectively. The signals coming from these exogenous tissues are essential for the initiation and progression of the specification of the optic vesicle into NR and RPE (Esteve and Bovolenta, 2006). I will now explain in more detail the regulation of NR and RPE specification genes by extraocular signals.

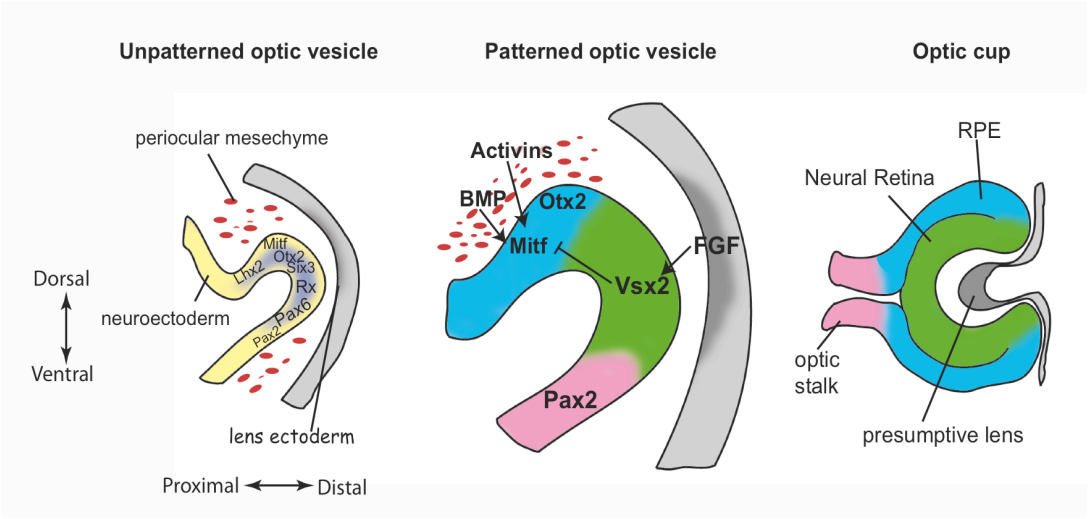


Figure 6. Optic vesicle patterning. In the unpatterned optic vesicle, neuroepithelial cells are indistinguishable and express a common set of transcription factors. In the patterned optic vesicle, activins and BMP signals from the periocular mesenchyme promote the RPE fate (blue), and FGF signals from the lens activate NR identity (green). *Vsx2* expression in the distal optic vesicle defines the NR identity by suppressing *Mitf*. At optic cup stage, the three tissues are already specified into RPE, NR and optic stalk (pink).

Step 3.1: The Retinal Pigmented Epithelium (RPE)

RPE specification is controlled fundamentally by the expression of two transcription

factors Mitf and *orthodenticle homeobox 2* (Otx2). Otx2 is required for *mitf* expression. Together they specify the RPE domain at optic vesicle stage and later on development they transactivate the expression of pigmentation genes. *mitf* and *otx2* are initially expressed in the eye field; their expression persists to optic vesicle stage and is eventually downregulated as the presumptive NR pattern emerge (Martinez-Morales et al., 2001; Martinez-Morales et al., 2004).

Signaling molecules secreted by the periocular mesenchyme are required for *mitf* expression during the early patterning of the RPE.

Experiments in chick embryos showed that the TGFb family member activin A, expressed in the extraocular mesenchyme, contributes to the specification of the RPE (Feijen et al., 1994). Chick optic vesicle explants cultured separated from the surrounding mesenchyme lose the expression of RPE markers and specific NR markers appear up-regulated (Fuhrmann et al., 2000). Moreover, these phenotypic defects are prevented by the addition of the TGFb family member activin A to the cultured cells (Fuhrmann et al., 2000), thus indicating that optic vesicle cells require signals from the extraocular mesenchyme to differentiate into RPE.

Bone Morphogenetic Proteins (BMP) BMP2, BMP4 and BMP7 are cell cycle regulators expressed in the periocular region and they also participate in the patterning of the optic vesicles (Dudley and Robertson, 1997). Previous studies in chick embryos showed that the overexpression of the BMP antagonist Noggin causes severe alterations in the ventral optic cup, including the loss of expression of RPE markers, which are then substituted by optic stalk markers (Adler and Belecky-Adams, 2002). These results suggest that the BMP signaling pathway expressed in the periocular mesenchyme is also essential for RPE specification (Figure 6).

Step 3.2: The Neural Retina

The presumptive NR is specified by the master gene *vsx2*, which is essential both for the patterning of this tissue at the optic vesicle stage and for the specification of NR cells at later stages. Vsx2 is the first TF expressed in the presumptive NR, in all undifferentiated retinal progenitor cells (RPCs). At this stage its required to maintain RPCs in a multipotent state (Vitorino et al., 2009). As development proceeds, the RPCs differentiate and *vsx2* expression remains only in bipolar interneurons and

Muller glia cells (Burmeister et al., 1996; Belecky-Adams et al., 1997; Chen and Cepko, 2000).

Mutations in *vsx2* are associated with microftalmia in humans (Bar-Yosef et al., 2004) and *vsx2* mutant mice (ocular retardation mutant - or^{J}) also show this phenotype (Truslove, 1962; Osipov and Vakhrusheva, 1983; Burmeister et al., 1996). In these mutants the retinal cells proliferation was decrease and the RPE differentiation genes were up regulated. Additionally, in *vsx2* knock-down zebrafish embryos, it was also observed malformations and morphological defects in early stages of eye development (Barabino et al., 1997). These findings indicate that Vsx2 is necessary for NR patterning.

During presumptive NR specification, signaling molecules derived from extraocular tissues regulate *vsx2* expression within the optic vesicles. Previous studies have shown that FGF signaling, emanating from the presumptive lens ectoderm, activates *vsx2* (Vogel-Hopker et al., 2000). The genetic interaction between FGF and *vsx2* is essential for retina patterning in the distal optic vesicle and, at later stages, for retinal neurogenesis (Rowan et al., 2004; Horsford et al., 2005). Furthermore, the conditional inactivation of *shp2* (which is required for complete activation of FGF singling) at the early optic vesicle stage in mouse mutants results in loss of *vsx2* expression. *Shp2* conditional mutants gain *mitf* expression in the NR; the affected part of the optic cup acquires RPE-like morphology and even gains some pigmentation (Cai et al., 2010). This finding suggests that FGF signaling regulates *vsx2* during NR specification, at optic vesicle stage (Figure 6).

In summary, the timely action of conserved transcription factors and inductive signals at optic vesicle stage, directs the initial development of the different compartments of the vertebrate eye. The regionalization of the optic vesicle is accompanied by the differential expression of transcription factors: Pax2 in the prospective optic stalk; Pax6, Rx, Lhx2 and Vsx2 in the prospective NR and Pax6, Otx2 and Mitf in the prospective RPE (Chow and Lang, 2001; Martinez-Morales et al., 2004; Bharti et al., 2006). After patterning has occurred, the optic vesicle undergoes specific morphogenetic movements that make the tissue invaginate forming the optic cup (Figure 4-C and D).

Step 4: Optic cup morphogenesis

The early organogenesis of the vertebrate retina finishes with the folding of the optic cup, leading to a perfectly shaped spherical eye. Geometrically, the invagination of the optic cup is a complex morphogenetic movement, which comprises different cell morphology changes. The invagination process has been examined using histology and life imaging in zebrafish embryos. It begins when the most distal part of the optic vesicle contacts the overlying surface ectoderm. This interaction results in the specification of the lens placodes, at the surface ectoderm, and precedes both tissues invagination (Li et al., 2000).

During invagination, the NR and RPE enwrap the lens as it emerges from the overlying ectoderm. The cells from the presumptive NR extend toward the forebrain, elongate and are subject to an anterior rotation (Kwan et al., 2012). The orientation of these cells at this time point is regulated by the combination of FGF signals along the dorsal-ventral axis of the neural tube (Picker et al., 1999). At the dorsal optic vesicle, the cells from the presumptive RPE follow the extension and elongation movements of the NR cells and after that, they spread dramatically to cover the back of the NR, becoming a flat monolayer where the distance between the cell nuclei is increased (Kwan et al., 2012).

Although there are detailed reports of the morphogenetic movements of the cells during optic cup folding (Li et al., 2000; Kwan et al., 2012), the evidences about the molecular players involved in the initiation and progression of this process started to be investigated only a few years ago.

Step 4.1: Lens ectoderm contributions to the optic cup folding

Previous studies in mouse showed that during the optic cup folding the cells from the surface ectoderm and the cells from the presumptive NR get in contact through F-actin-rich basal filopodia. Basal filopodia are cytoplasmatic projections of the basal membrane of the cells. It has been suggested that the interaction between the filopodia from neuroepithelial cells and from the lens ectoderm act, as a fine-tunning mechanism, to assist the mechanical forces necessary to coordinate the folding of the optic vesicle and the presumptive lens (Chauhan et al., 2009).

Although these findings indicate that the physical interactions between the lens

ectoderm and the optic vesicles may play a role during optic cup folding, several studies demonstrate that this process can also occur without physical contacts from this adjacent tissue. The interaction between retina and lens ectoderm surface also activates signaling pathways that might contribute to the folding of the optic cup facilitating its morphogenesis (Fuhrmann, 2010). Previous studies in chick embryos showed that in the absence of pre-placodal ectoderm (early optic vesicle stage), the invagination of the optic cup is perturbed. However if surface ectoderm ablation occurs later, the optic cup invaginates properly, even without a correct formation of the lens (Hyer et al., 2003; Smith et al., 2009). This result suggests that essential signaling molecules for optic cup folding are expressed in the lens ectoderm at early stages, at the onset of optic vesicle patterning.

More recently, *in vitro* work using mouse embryonic stem cells showed that isolated cell aggregates are able to organize spontaneously into a retina. These cells form a neuroepithelium and bend into hemispherical vesicles that became patterned with retina marker genes along their proximal-distal axis. This tissue could then fold on its own, without interaction with exogenous tissues and in a very similar manner to the folding of the retinal epithelium in the embryo. This work therefore suggests that the optic vesicle contains self-organizing programs, which involve stepwise and domain-specific regulation of the different optic vesicle domains. These programs are responsible for the morphological events and cell identity during optic cup folding (Eiraku et al., 2011).

Step 4.2: Coordination of cell shape changes during optic cup morphogenesis

The morphogenetic events forming the optic cup require the generation of tensional forces to coordinate cell shape changes and movements. Morphogenetic events can occur in animal epithelial sheets either by apical or basal constriction of the cell membranes (e.g. Sawyer et al 2010, Gutzman et al., 2008). Both processes require the recruitment of the actomyosin network, which is responsible for cell contraction (He et al., 2010). Previous work in *Drosophila* embryos showed that apical constriction depends on apical polarity complexes and adherens junctions between cells (Kolsch et al., 2007; Letizia et al., 2011). Basal constriction, which has been studied in *Drosophila* and medaka embryos, depends on integrin-mediated focal adhesions.

Interference with the adhesive function of integrins impairs basal actomyosin recruitment and tissue morphogenesis in both vertebrate and invertebrate epithelia (Martinez-Morales et al., 2009; He et al., 2010). Integrins are well-characterized transmembrane receptor proteins that mediate the attachment between a cell and its surroundings (Hynes, 1992). These proteins, localized within focal adhesion complexes, mediate the regulatory effects of extracellular matrix adhesion on cell behavior, like mechanical forces and signaling transduction (Burridge and Chrzanowska-Wodnicka, 1996).

Previous studies in medaka demonstrate that optic cup folding occurs through basal constriction (Figure 7). This work shows that optic cup folding requires the adhesive function of integrins, which is dependent on the expression of the developmental gene *opo*. During the optic cup folding Opo regulates the localization of the integrins promoting basal constriction of the cells and optic cup folding (Martinez-Morales et al., 2009).

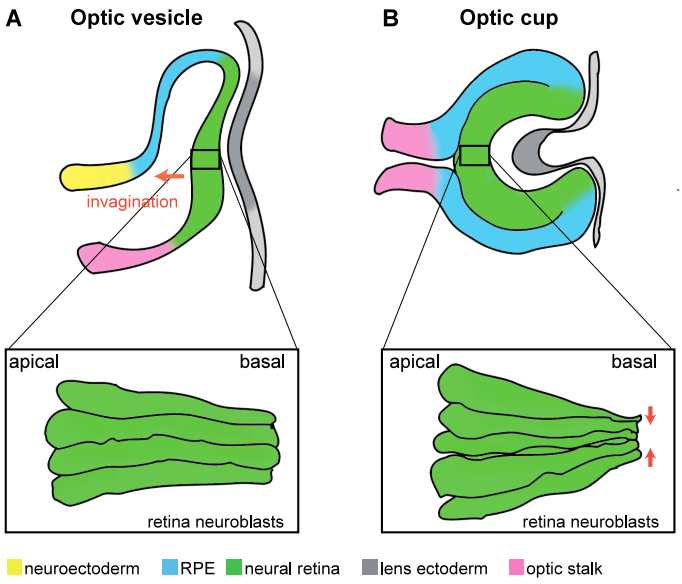


Figure 7. Optic cup folding

A) At optic vesicle stages, retina neuroblasts are organized in a monolayer at its distal part.

B) During the invagination of the optic cup, the basal membrane of the cells is constricted (red arrows) to force the optic cup folding.

4. *Opo* is required for optic cup folding

opo was identified in medaka in a large-scale screening using ENU-mutagenesis (*N*-ethyl-*N*-nitrosourea) as a recessive lethal mutation with full penetrance and low phenotypic variability (Loosli et al., 2004). In *opo* mutants, eye morphogenesis defects became evident during the optic cup folding stage. The expected basal

surface contractions are not observed in the retina neuroblasts, which indicate that tension cannot be exerted properly when *opo* function is damage. Thus, as development proceeds, the optic cup does not fold properly and therefore, the mutants show abnormal morphologies, with large ventral openings in the eyes (Martinez-Morales et al., 2009).

Opo is a developmental gene placed in a gene desert enriched with cis-regulatoty elements. It encodes for a transmembrane protein, localized in the compartments of the secretory pathway and basal end-feet of the neuroepithelial precursors cells. During optic cup development, this protein regulates the asymmetric localization of integrins to the basal surface of the retina epithelium (Martinez-Morales et al., 2009).

The trafficking of integrins along the apical basal axis has been proven to be essential for integrins adhesive function, basal actomyosin recruitment and further a directional cell movement (Ezratty et al., 2009). The endocytosis of integrins is a clathrin-dependent process regulated by members of the phosphotyrosine binding (PTB) family of clathrin adaptors (such as Numb, Dab2 and ARH), which interact directly with integrins and regulate their internalization (Calderwood et al., 2003). This interaction has been shown to provide a directional cell migration in HeLa cells (Nishimura and Kaibuchi, 2007).

Opo functions as a repressor of Numb and hence interferes with clathrin-mediated integrin endocytosis. This antagonism between *opo* and *numb* controls focal adhesions turnover and is specifically required for optic cup folding (Bogdanovic et al., 2012a). Accordingly, in *opo* mutants the lack of adhesive function (i.e. due to an excesive integrin internalization) impairs basal actomyosin recruitment and subsequently optic cup folding.

These findings indicate that the basal constriction mechanism that directs the optic cup folding depends on the molecular antagonism found between Opo and Numb. Future studies will need to address what kind of signals initiate invagination and how conserved is the mechanism mediated by *opo* across vertebrate species.

5. *Opo* role in the development of other tissues.

Besides the optic cup defects, *opo* mutants also show defects in other epithelial tissues like brain, heart, fins and in neural crest-derived craniofacial structures

(Martinez-Morales et al., 2009).

The development of the craniofacial region mostly depends on the multipotency and migratory behavior of neural crest cells. This cell population is considered a vertebrate innovation and, accordingly, chordate ancestors lacked neural crest counterparts (Santagati and Rijli, 2003).

In *opo* mutants neural crest cells are correctly specified, however, they fail to migrate/delaminate, as is shown by the severe reduction in the jaw and palate structures. As is the case for the neural crest precursors, fin bud mesenchyme is also correctly specified in *opo* mutants, however the apical ectodermal ridge is misshapen. As a result, pectoral fins fail to grow properly (Martinez-Morales et al., 2009).

Analysis of the *opo* expression pattern in mouse has shown a similar expression pattern to that described in medaka. This includes craniofacial neural crest derivatives, skeletal elements, heart and the eyes (Santagati and Rijli, 2003; Mertes et al., 2009).

In terms of protein conservation, the Opo carboxy-terminal motif is homologous to proteins present in basal chordates, echinoderms, cnidarians (not in protostomes) and in the metazoan basal lineage. The N-terminal motif of the protein can be considered a vertebrate innovation. Taking into account that *opo* has an important role in the development of organs such as the neural crest, the fins or the chamber-shape eye, which are all considered vertebrate novelties, it is likely that *opo* function co-evolved with the emergence of these vertebrate tissues (Shimeld and Holland, 2000; Martinez-Morales et al., 2007; Martinez-Morales et al., 2009).

In summary, *opo* functions as an integrin endocytosis regulator at the basal surface of the retina, during optic cup folding. Interference with its activity impairs optic cup morphogenesis. Given the crucial role that *ojoplano* plays during eye and neural crest development, and the intriguing link that it appears to have with human hereditary diseases, we decided to analyze the enhancers that control its expression. The localization and analysis of these elements allows the investigation of upstream regulators of *opo*. Furthermore, the identification of these enhancers may establish new links between *opo* and the TF network implicated in early eye development. In addition, it was a challenge to link these new enhancers with *opo* associated hereditary diseases (see Box 1).

To achieve this, I have identified a collection of enhancers within *opo locus* by enhancer screening, using zebrafish as model organism (Bessa and Gomez-Skarmeta, 2011). I generated a collection of stable zebrafish transgenic lines for each enhancer, which were used to characterize its spatio-temporal activity and its potential upstream regulators.

II. Materials & Methods

1. Selection of non-coding putative regulatory elements

The selection of the highly conserved putative *non*-coding regulatory elements (n=23 conserved PNREs), was carried out using the VISTA Browser (http://pipeline.lbl.gov/cgi-bin/gateway2; (Frazer et al., 2004), taking as a reference the human genome released in March 2006 (NCBI36/hg18) and using the default parameters (regions of al least 100 base pairs (bp) and 70% similarity.

To localize the epigenetic marks H3K4me1 and H3K27ac in the human PNREs, ChIP-seq tracks from six human cell lines were used: embryonic stem cells (H1hesc), mammary epithelial cells (Hmec), skeletal muscle myocites (Hsmm), umbilical vein endothelial cells (Huvec), epidermal keratinocytes (Nhek) and lungs fibroblasts (Nhlf) data (available in UCSC browser: http://genome.ucsc.edu/; (Bernstein et al., 2006; Mikkelsen et al., 2007). Four PNREs showing no conservation between vertebrates, but having high H3K4me1 and H3K27ac signal in human cell lines, were also selected. The PNRES were named according to the genome: human genome (H), chromosome 6 (6), and the first five digits of the element localization (i.e.: the peak localized in chr6:9040802-9041496 is named: *H6:09040;* Supp. Tab. 1).

To investigate if the conserved PNREs were marked with H3K4me1 and H3K27ac signatures during zebrafish developmental stages, previously published ChIP-seq data was used (Bogdanovic et al., 2012b) (Supp. Tab. 2).

2. Cloning into a Tol2-mediated vector

The PNREs were amplified by PCR from human genomic DNA using the primers listed in table 1.

Name	Forward primer (5'-3')	Reverse primer (5'-3')	Localization in human genome (hg18, chromosome 6)		Lenght (bp)
			start	end	
H6:09040	CCATATGGATCATCCTCTTCCC	CTGGAAGCGGTTGTTTGAGG	9040802	9041495	886
H6:09360	TCCTTTCTCCCCAAAACCAC	TTTGTGAATAATGTTTGCACTAAGAAG	9360754	9362917	2360
H6:09510	TTTGAGCTTGCTGCATTATATGG	AATTTGCTTTGGCAGTTGGG	9510079	9512442	2656
H6:09516	GGGGGAAATCTATGTGCATTTG	ACTGTTCATGTGGCTGGGG	9516753	9517809	1216
H6:09536	GCTCAAGGAAGCTGTGATGTC	TCTGTGAATGTGCATTTGTTCG	9536004	9537930	2027
H6:09563	CACAAGGCATCTTACACATCAGC	GCGATCAGGCGGTCTAAATC	9563974	9565999	2148
H6:09601	ATTGCCACATAATGAGTTGCAGTCC	CTGGCCACACATCACAAGTCAAAG	9601172	9602202	1031
H6:09617	GGATGTGCTCCAAGTCACCAAG	AGTGCCAGCAGAGACCATTC	9617863	9618922	1060
H6:09628	CATCTGCTTGGCTTCTGATG	ATGGCTATGGGGGAGATAGG	9628639	9629494	856
H6:09647	GGAGGAGAGAGAGCAGCGAG	AAATGAACTGCAGCCAAGATG	9647690	9648627	1121
H6:09671	TGCTTCATCCCGTAGGTTTGTTTC	AAAGGGGTATGGGAAAAGGTGTCAG	9671795	9672798	1004
H6:09742	CAAACAAGGATTTCTCAAGGGC	CATTAGTCATGATTGAAAAATGGATG	9742877	9745652	2856
H6:09778	AATTGCTTATGGTTGTTCCCG	TCAAAACCCTAACTTGAAGCAAAAC	9778300	9780223	2186
H6:09813	GATCTTCAACATGACTTCTGGC	TTGGGTTAGGTCCTTCAACTGTG	9813174	9814241	1068
H6:09815	GCCCATAAGGTGAGACTATAACTG	ACATCTTGGCTATTGTGAATAATGC	9815794	9816729	936
H6:09892	GTGCTCCAATGCAGTTGCTC	TTAATATCCCAATTTGATACTTGGC	9892240	9894126	2144
H6:09941	TCATGTCTTCCAATGAGTCTGG	GCAAAAAGAAAGAGCCGAGA	9941162	9943190	2029
H6:09949	GGCAACCCAAGTTAGTAGCACCACC	TCAGTAAGGCCAGGCTCCATC	9949629	9950467	839
H6:10043	AGCCTGGGTAACAGAGCGAG	GCATTCAGCCCTACCCAGTC	10043446	10045884	2571
H6:10126	CTGATTTCAAATGCATGCCC	AAATTCCAGAGGGCAGATGG	10126474	10127046	685
H6:10137	CCAAGAAGAGACCTCCAACGAAC	CACTTTTATGAGCTACGTGGTTGTG	10137953	10139056	1104
H6:10258	CTGACAGCTCGCATCGTCCAC	GAGTATAAGGTGAGGTTGGCAAC	10258016	10259047	1032
H6:10309	GTTGCGATGTGTTGTTTTGG	GTCCAGTGGTGGCAGACTA	10309982	10311174	1193
H6:10377	CACATCACAGCCAGAGAAGGGTC	TGGCACGTGCTTACCTATGTAAC	10377252	10378117	866
H6:10422	AGCTTTTCTCCGCTTTGCTG	TTCTTCCCTCCTTCTTCTGGTG	10422720	10424962	2511
H6:10431	ACCTTCCTTTGTGCTTGCCATCC	CTACTGCACCCGACCCTGACTC	10431893	10433099	1207
H6:10455	TAGCACCATTGCACTTCAGC	CTTCTGCTTCATTTGCCACA	10455468	10456853	996

Table 1. Primers used for PNREs amplification.

The PCR fragments were sub-cloned in PCR8/GW/TOPO (Invitrogen) vector and, using the Gateway Technology (Invtrogen, CA, USA), they were then transferred to the destination enhanced green fluorescent protein (eGFP) reporter vector, as described by Bessa and colleagues (Figure 9; (Bessa et al., 2009).

Figure 8. The Zebrafish enhancer detector (ZED) vector. Contains two modules flanked by Medaka Tol2 transposase target sites (orange), that enables efficient transgenesis in zebrafish. The first module contains the minimal GATA promoter (light blue) driving the expression of eGFP (green). Each PNRE was cloned upstream of this module using the Gateway Technology recombination (Gateway *cassette* in yellow), activating or not the GFP expression. Two strong insulators are present in the ends of this first module to reduce the potential influences of the regulatory elements (dark pink). The second module contains the *cardiac actin* promoter (light blue) coupled to a red fluorescent protein (RFP in red), which serves as a positive control for transgenesis. The positive embryos should express homogeneous RFP in somites and heart (*adapted from Bessa et al., 2009*).

3. Maintenance of zebrafish (*Danio rerio*)

Zebrafish was used as a model organism, given their characteristic capacity to deliver large amounts of eggs *per* week, the transparency of their newborn embryos and its fast organogenesis (approximately 3 days). These characteristics are an advantage to investigate morphogenesis and organogenesis during early development.

The wild-type AB/Tuebingen (AB/TU) zebrafish strain was maintained under standard conditions according to the procedures described in Kimmel et al. (1995) and in the Zebrafish Model Organism Database (http://zfin.org; (Sprague et al., 2003). All developmental stages in this study are reported in hours post-fertilization (hpf) at 28.5°C in E3 medium with 0.003% 1-phenyl-2-thiourea, which prevents pigmentation (Figure 9) (Kimmel et al., 1995).

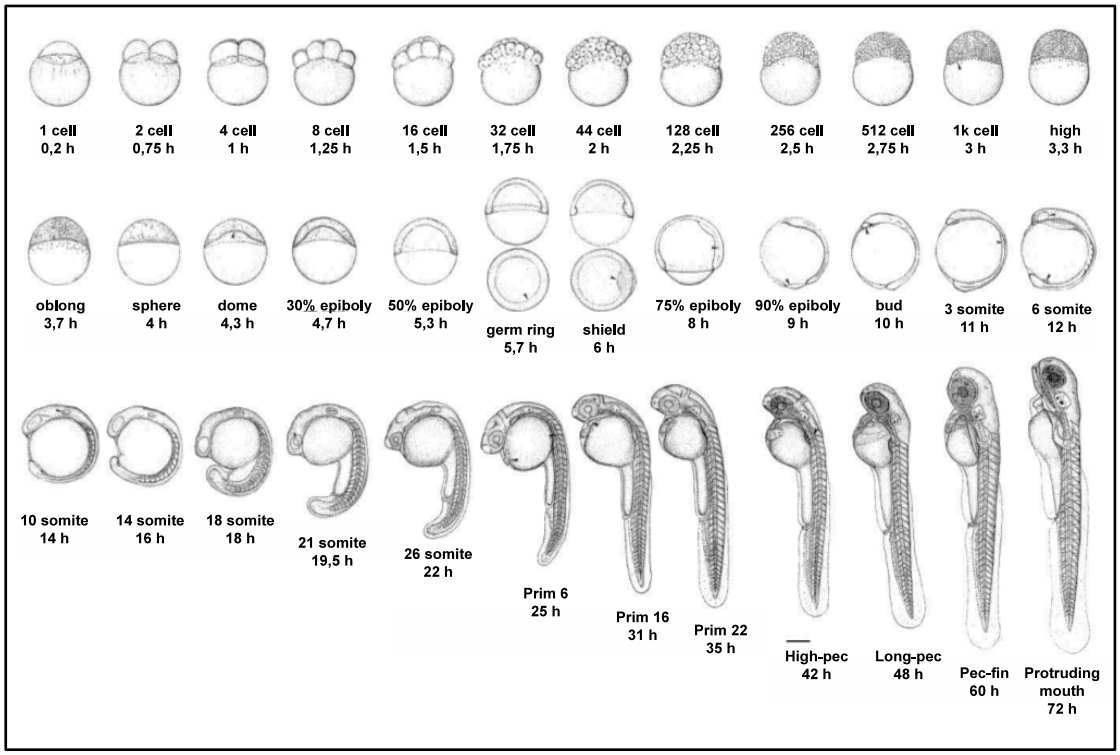


Figure 9. Developmental stages of zebrafish embryo. From one cell stage to 72hpf (*adapted from Kimmel et al., 1995*).

4. Zebrafish Tol2 mediated transgenic assays

The ZED constructs containing the PNREs were microinjected in zebrafish embryos using the following procedures. In the day previous to DNA microinjection, the adult fish were separated according to sex into two different 3 litter (L) tanks. By the

following morning, males and females were placed together in a 5L tank, provided with a grid to separate the adults from the newborn eggs (Kimmel et al., 1995). Minimums of 300 fertilized eggs were injected, at one cell stage, with 3-5 nl of a solution containing 25 nM of PNRE construct, 25 nM of Tol2 mRNA and phenol-red (Sigma-Aldrich) at a final concentration of 0.1%. The microinjected embryos where then incubated at 28°C, during 3 days, and an equal number of un-injected embryos were also incubated as experimental control.

The GFP and RFP patterns of the injected embryos were observed at 24, 48 and 72hpf with a fluorescent stereoscope. Embryos showing homogenous and strong RFP patterns in somites and heart (putative transgenic embryos) were raised to sexual maturity.

5. Tissue specific enhancer screening in F1 generation

Putative transgenic adult fishes (founders) were individually outcrossed with a wild-type zebrafish partners. At least 10 individual fishes were analysed *per* PNRE. The GFP and RFP expression patterns of the first generation embryos (F_1) were observed at 24, 48 and 72 hpf. A PNRE was considered a tissue specific enhancer when three or more founders displayed similar F_1 GFP patterns during development (Bessa et al., 2009). To obtain high-resolution pictures, a fluorescent stereoscope set up with a digital camera coupled was used. Adobe Photoshop was employed to adjust brightness and contrast of the obtained images. Finally, the expression pattern of the new enhancers was compared to the *opo* expression pattern obtained from RNA *in situ* hybridizations and the expression analysis of the *opo* promoter region.

6. Wholemount RNA *in situ* hibridizations and Fluorescent RNA *in situ* hibridizations (ISH & FISH)

To synthesize *opo* riboprobes, total RNA was extracted from wild-type zebrafish strain at 24hpf using Trizol (TRI Reagent: Acid-guanidine thiocyanate, Sigma) followed by phenol-chloroform extraction. Genomic DNA was then digested using TURBO Dnase (DNAseI, *Boehringer Mannheim*) and the mRNA was precipitated with 7.5M of LiCl. Around 1–5 µg/µl of the total RNA was reversely transcribed using a cDNA synthesis Kit (*Invitrogen*). The cDNA was used as template to amplify *opo*

fragments using the following PCR primiers: 5' AAGTTGCAGCAGAAGGCTGTGGAG and 5' ATTTTGCGGAATTGGTTGAG for N-terminal; 5' GTCAGCTAGTGTTTGGCTCAC and 5' ATGCATTAGTGTCAGAATCAT for C-terminal probe. The PCR products were cloned into pCS2+ vector (Addgene), linearized with *NotI (Takara Bio Company)*, and transcribed with SP6 (*Roche Applied Science*) using digoxigenin-labelled nucleotides (Digoxigenin-11-UTP, *Roche*). The riboprobes were used at a concentration of 1.5 ng/ul for ISH and FISH. ISH and FISH were preformed as described in (Neto et al., 2012) and (Bogdanovic et al., 2012a) respectively. Embryos processed for ISH were observed with a scope (Leica) and FISH embryos observed with a confocal microscope (Leica SPE). The ISH embryos were then processed for vibratome sectioning, as described in (Martinez-Morales et al., 2009).

7. Cloning of the medaka *opo* promoter region

The genomic DNA was extracted from medaka embryos (*Oryzias latipes*) with phenol:chloroform:Isoamyl alcohol 25:24:1 (Sigma) and washed with 100% ethanol. This DNA was used as a template to amplify a 5,6kb region upstream of *opo* transcription start site (defined using RACE experiments in Martinez-Morales et. al. 2009) with the following primers: 5' GTCAGCTATCACCGCAGTCGTCA and 5' CCTCACTTCCTTGTGCCGACATG 3'. After purification, the 5,6kb PCR product was cloned in PCR8/GW/TOPO vector (Invitrogen), transferred into the destination ZED vector (Bessa et al., 2009) and used for transgenesis in zebrafish as above mentioned, in the sections 3, 4 and 5.

8. Circularized Chromosome Conformation Capture Analysis (4c-seq)

4C-seq assays were performed as described in (Dekker et al., 2002; Hagege et al., 2007; Noordermeer et al., 2011; Splinter et al., 2012). Newborn mice at stage E11.5 were dissociated by collagenase and fixed in 4% formaldehyde. The sample was then treated with lysis buffer (10 mM Tris-HCl pH 8, 10 mM NaCl, 0.3% IGEPAL CA-630 (Sigma), 1X protease inhibitor cocktail (Roche). Nuclei were digested with DpnII endonuclease (New England Biolabs) and ligated with T4 DNA ligase (Promega). Subsequently, Csp6I endonuclease (Fermentas, Thermo Scientific) was used in a second round of digestion, and the DNA was again ligated. Specific primers were

designed near *opo* promoter and H6:10137 enhancer (viewpoints), with the online program primer3 v. 0.4.0 (Rozen and Skaletsky, 2000). Illumina adaptors were included in the primers sequence. The PCRs were performed with Expand Long Template PCR System (Roche) for each viewpoint, pulled together and purified using High Pure PCR Product Purification Kit (Roche). The samples were then quantified using Quanti-iT^{TM} PicoGreen dsDNA Assay Kit (Invitrogen) and submitted to deep sequencing. The analyses of the 4C-seq were performed as described in (Noordermeer et al., 2011). Briefly, raw sequencing data were demultiplexed and aligned using mouse genome released in July 2007 (NCBI37/mm9) as reference. Reads located in fragments flanked by two restriction sites of the same enzyme, or in fragments smaller than 40 bp were filtered out. Mapped reads were then converted to reads-per-first-enzyme-fragment-end units, and smoothed using a mean running window algorithm. To calculate the statistically significant targets for each viewpoint, the average background level was estimated randomizing data in a window of 2 Mb around each viewpoint. The *p*-value for each potential target was calculated by means of Poisson probability function. Smoothed data were uploaded to the UCSC Browser for visualization.

9. Search for up-stream regulators of H6:10137 enhancer

To search for conserved transcription factors DNA-binding sites in H6:10137 sequence it was used the *Mulan* alignments and the *multiTF* search tools were used, which are available in ECR browser, which detects DNA-binding sites based on sequence conservation (TRANSFAC 7.0 *in* http://ecrbrowser.dcode.org/; (Ovcharenko et al., 2004). In this case the binding sites were predicted with a sequence homology of 63% between human and *Xenopus tropicalis.* Additionally it was used the JASPAR database (http://jaspar.genereg.net; (Bryne et al., 2008). JASPAR database contains a collection of DNA-binding sites for transcription factor, modeled as matrices that are compared with a database of experimentally defined transcription factor binding sites (Stormo, 2000; Wasserman and Sandelin, 2004). For JASPAR data, motifs with a sequence homology of 100% and a socre >10 (maximum score=13,5) were considered.
The expression pattern of the founded motifs in ECR and JASPAR was analyzed in the Zebrafish Model Organism Database (http://zfin.org/; (Sprague et al., 2003)). The transcription factors that shared a common expression pattern with the H6:10137

enhancer were then analyzed with multi-species alignments using *ClustalX 2.0* (Larkin et al., 2007).

10. Mutagenesis of Vsx2 DNA-binding sites inside H6:10137 sequence

The Vsx2 DNA-binding sites in the H6:10137-ZED construct were mutated by PCR, using QuikChange Multi Site-Directed Mutagenesis Kit (*Agilent technologies*) with the mutagenic primers: 3'CTATTTTACATTTTCTTAATcgcCAAATCTTAATCTTGACTTTAATG'5 (mutH6:10137del) and 3'CATTATTAGCGGTCTctgCTTGAAGTTATATCAATGTTAGT'5 (mutH6:10137del1), following the manufactures protocol. After sequence disruption confirmation by sequencing, the mutagenic-H6:10137 construct (mutH6:10137) together with the H6:10137 enhancer construct were injected in wild-type zebrafish at one cell stage, as described before. The GFP and RFP activity of both, mutH6:10137 and H6:10137, were compared in F0 (mosaic embryos) at 24hpf.

11. *Vsx2* Morpolinos injections

Vsx2 splicing morpholinos (*vsx2*MOs) were obtained from Gene Tools: I1E2MO: 5'-TCATCTGAATCTGTAACATGGAGGA-3' blocks a splicing acceptor between Intron 1 and Exon 2, E2I2MO: 5'-TTTTCTTTAAGCCCACCTGTGTCGT-3' blocks a splicing donor between Exon 2 and Intron 2.

Both morpholinos were separately injected at one-cell stage in wild-type strain and in the transgenic lines: Tg(H6:10137) and Tg(*vsx3*GFP:caax), in a concentration of 8,5-13ng per embryo. When co-injected in a wild-type strain, the concentrations used were 6-10ng for each *vsx2*MO. *P53* morpholino (*p53*MO: 5-GCGCCATTGCTTTGCAAGAATTG-3 (Langheinrich et al., 2002), was co-injected with *vsx2*MOs in a concentration of 4-6ng per embryo. Control embryos were injected in parallel with water and *p53*MO alone. To assess the morpholino efficiency, cDNA samples were obtained from 24hpf sectioned heads and total embryos, from both, controls and morphants embryos, using Trizol for the mRNA extraction followed by cDNA synthesis, as described before (section 6). The cDNA samples from both morphant and control embryos were also used to quantify *opo* expression levels quantification by RT-qPCR (see section below).

The amplification of the *vsx2* transcripts in both morphant and control cDNAs was done by RT-PCR using the following primers: I1E2MOfw: 5'-AGCACACTGGACTCCTTCCCCG-3'; E2I2fw 5'- GGGATTAATTGGGCCTGGAGGTAT-3'; I1E2MOrv and E2I2rv: 5'- AGACTCGGGCAGAGGGATAGAGTGA-3'. The PCR products were cloned into StrataClone® vector (Agilent) and sequenced (Secugen) to confirm aberrant splicing events induced by the morpholino.

12. *Opo* RT-qPCRs

The quantification of *opo* expression levels in zebrafish cDNAs by RT-qPCR was done using a forward primer placed in second exon: 5'CGGTCTTCGCTGTAGATGC'3 and a reverse primer placed in third exon: 5'GACGGCGTAGAGGAATAATG'3. The results were normalized with *eif1a*, as a *house-keeping* gene, amplified with 5'-CTTCTCAGGCTGACTGTGC-3' as forward primer and 5'-CCGCTAGCATTACCCTCC-3' as reverse primer.

13. *Vsx2*MOs - Eye phenotype characterization

The eyes of both, *vsx2*MOs injected embryos and its respective controls were observed at 24hpf, using a fluorescent scope coupled with a photo camera (SZX16–DP71, *Olympus*). Photos were taken with the same parameters of magnification, sensitivity and exposition in both transmitted light and fluorescent conditions. The GFP quantifications, the eye ventral openings angles, the retina length and the forebrain height were measured for each embryo with *ImageJ1.45q* (Wayne Rasband, National Institute of Health, USA). The retina length values were normalized with the forebrain size for each individual embryo. Over 100 embryos, including their respective controls were analyzed for each phenotypic condition.

14. *Vsx2* Overexpression

The total RNA extraction and cDNA conversion from zebrafish wild-type embryos were preformed as above mentioned in section 6. The *vsx2* cDNA was amplified by PCR using the primers 5'-GGACCGACTCGAAAGCAAAG-3' and 5'-TTAACGCACAAGCCCAAGTC-3'. The PCR products were cloned into pCS2+ vector

(Addgene), generating the Vsx2mRNA:pCS2+ construct. This construct was linearized with *NotI (Takara Bio Company*), and transcribed with mMESSAGE mMACHINE SP6 Kit (*Ambion*), following the manufacture protocol. After the capped mRNA transcription, the DNA template was eliminated. *Vsx2* mRNA was precipitated with 7.5M LiCl and quantified. The injections of the *Vsx2*mRNA were preformed in zebrafish wild-type at one cell stage (200pg and 340pg *per* embryo). At 24hpf the injected embryos were processed for total RNA extraction followed by cDNA synthesis as mentioned before. The cDNA was used for quantification of *opo* expression levels, by RT-qPCR, as previously described.

15. Chromatin Immuno-precipitation detected by qPCR (ChIP-qPCR)

The physical interaction between H6:10137 and Vsx2 was tested by ChIP-qPCR using the fusion protein *GFPvsx2*mRNA:pCS2+ and the GFP antibodies (Abcam). Briefly: wild-type zebrafish embryos were injected into one cell stage, with 300pg/embryo of *GFPvsx2*mRNA fusion. At 24hpf, 300 heads were dissected, from the injected embryos and fixed in 1% formaldehyde for 15min at room temperature. Following crosslinking, tissues were washed and sonicated (DNA to lengths between 200-300 bp). Half of the samples were stored at -20°C as input and the rest was immuno-precipitated overnight with pre-bounded antiGFP-Dynabeads (Sigma). Then, the immuno-precipitated DNA fragments and the Input sample were subjected to DNA extraction with phenol:chloroform:Isoamyl alcohol 25:24:1 (Sigma) and washed with 100% ethanol. RT-qPCRs were preformed using the following primers: 5´-CAGTGAAGAAGGAAAGCATGG -3´and 5´- TGCATCTTACAGAAGAAGAGA -3´ for H6:10137 enhancer; 5´-GTACCGATTTAATACTGTGG-3´ and 5´-CAGTATAGCATACGCTTTTA-3' for *Vsx1* (positive control obtained from: (Clark et al., 2008)); 5´-CTCTCAAATCCGGAGATGCT-3´ and 5´-GAAGCTCTCCACACACATGG-3´for *Ef1a* (housekeeping gene used as negative control). The recovery of the H6:10137 sequence from the immuno-precipitated sample was compared to the recovery of the positive and negative control sequences, taking the input as control sample.

III. Results

1. Enhancer screening in the *opo locus*

To explore the regulatory logic within the *opo* landscape, putative *non*-coding *cis*-regulatory elements (PNREs) were identified on the bases of their conservation in vertebrate genomes as detected by phylogenetic footprinting. The comparison between the genomes of mammals and teleosts allowed the identification of 23 PNREs in the *opo locus*. These elements were selected in the human genome and have a minimum of 70% of sequence homology between *Fugu* and human (Figure 10).

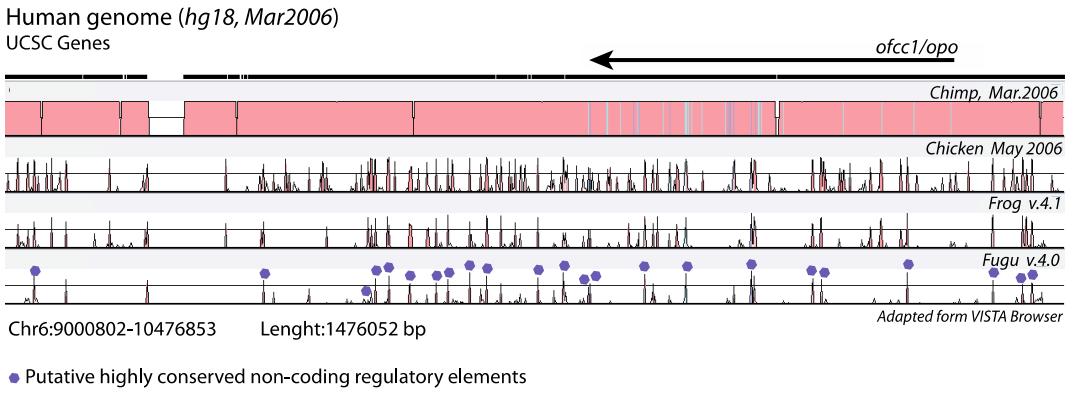


Figure 10. *opo regulatory* landscape. Highly conserved non-coding elements in the *opo locus* are depicted in the figure (purple circles marked in *Fugu* genome).

The majority of the conserved PNREs revealed the presence of typical epigenetic marks for constitutive (H3K4me1) and active (H3K4me1 and H3K27ac) enhancers, when compared to previously published data from chip-seq analyses in human cell lines (Bernstein et al., 2006; Mikkelsen et al., 2007) and zebrafish embryogenesis (Bogdanovic et al., 2012b) (Supp. Tab. 1-2). In 48% (n=11) of the PNREs, H3K4me1 and H3K27ac marks were observed in at least one human cell type and/or during zebrafish embryogenesis; In 43% (n=10) of the PNREs only H3K4me1 was observed in at least one human cell line and/or during zebrafish development and only 9% (n=2) did not shown any epigenetic mark for active or constitutive enhancers (Figure 11).

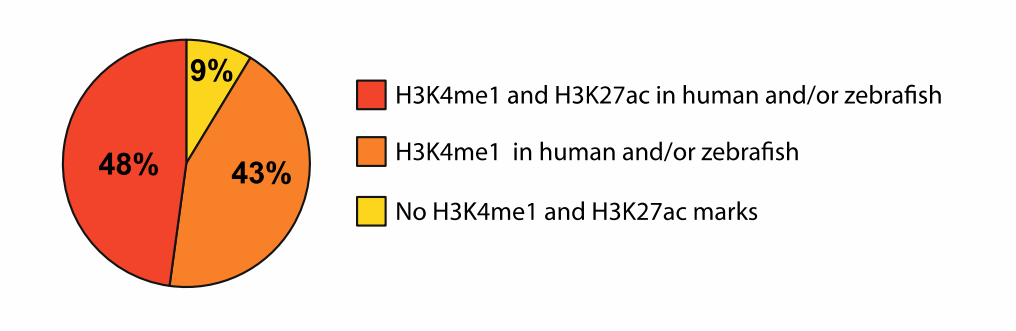


Figure 11. PNREs classification. Classification of the conserved PNREs according to epigenetic marks for enhancer activity in human cell lines and in the developing zebrafish embryo (24hpf and 48hpf).

In addition to the conserved PNREs, other 4 H3K4me1 and H3K27ac positive regions showing no conservation between human and teleosts were selected in the human genome as a reference (Supp. Fig. 1).

The selected elements were analyzed by transgenesis assays in zebrafish using the ZED vector (Bessa et al., 2009), a Tol2 mediated transgenesis shuttle that allows the integration of the humans PNREs in the zebrafish genome. The progeny (F1) of the putative transgenic fish (founders) was screened for reproducible GFP expression patterns between different founders. We observed that 44% (n=12) of the elements were able to activate GFP expression. These elements were all conserved between vertebrates (Supp. Tab. 3). However, from these 12 elements, only 9 behaved like tissue specific enhancers (n=9), meaning that three or more founders showed a reproducible expression pattern in their progeny (Table 2).

GFP + PNREs	Localization in human genome (hg18, chr:6)		n.º of Foundres			Classification
	start	end	Reproducible GFP pattern	Non-reproducible GFP patterns	Total	
H6:09040	9040802	9041495	3	-	13	Enhancer
H6:09516	9516753	9517809	4	-	23	Enhancer
H6:09563	9563974	9565999	8	-	25	Enhancer
H6:09601	9601172	9602202	3	-	13	Enhancer
H6:09617	9617863	9618922	3	1	13	Enhancer
H6:09778	9813174	9814241	-	5	16	Ambiguous
H6:09813	9815794	9816729	-	2	20	Ambiguous
H6:09949	9949629	9950467	3	-	10	Enhancer
H6:10126	10126474	10127046	-	4	20	Ambiguous
H6:10137	10137953	10139056	6	-	12	Enhancer
H6:10377	10377252	10378117	10	3	41	Enhancer
H6:10431	10431893	10433099	4	-	13	Enhancer

Table 2. PNREs with GFP activity. Classification according to the presence of reproducible GFP patterns between founders. If a PNRE shows similar GFP patterns in 3 or more different founders progeny, it is classified as a tissue-specific enhancer (enhancer). If the GFP patterns are different between the founder's progenies for one PNRE, this PNRE is classified as "Ambiguous".

The nine tissue specific enhancers were equally distributed at 5' and 3' from the *opo* sequence. They all showed sequence conservation from human to teleosts, and the majority (7/9) had both epigenetic marks for active enhancers in human cell lines and in zebrafish whole embryos (Figure 12).

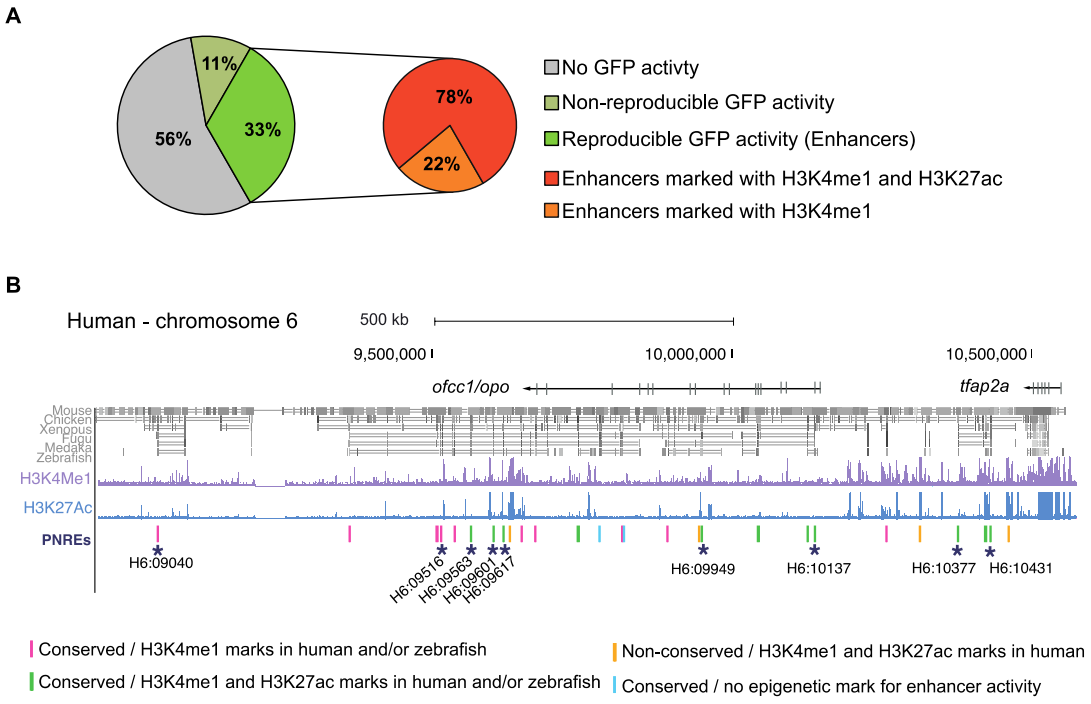


Figure 12. Summary of the enhancer screening within the *opo locus*. A) Percentage of PNREs acting as tissue-specific enhancers B) The *opo* regulatory landscape in the human genome (hg18 - chr6: 8936346-10570697, adapted from UCSC Genome Browser), showing sequence conservation from mammals to teleosts. The sum of H3K4me1 and H3K27ac marks in six human cell types is shown in purple and blue respectively; the last row represents the PNREs collection classified by color, with tissue specific enhancers identified and marked with an asterisk.

The analysis of PNREs in the *opo* landscape revealed that this locus is enriched in conserved non-coding elements that might be regulating the expression of the nearby genes.

2. Expression patterns of the *opo* locus enhancers

The GFP expression patterns of the nine enhancers found in the human *opo locus* were analyzed spatially and temporally, during zebrafish development. The enhancers did not display redundant expression patterns. However, they share common GFP patterns in a number of tissues. The enhancer patterns were distributed mostly within central nervous system (CNS), and craniofacial structures. A full description of them is shown in Table 3 and Figure 13.

***opo locus* Enhancers**	**GFP expression at early stages (16 to 36hpf)**	**GFP expression at late stages (36 to 72hpf)**	**Enhancer regulatory features**
H6:09040	CNS: more intense in the MHB	weak expression	5' from opo; H3K4me1;
H6:09516	FB and BA	FB, PL, BA ears and Jaw	5' from opo; H3K4me1;
H6:09563	CNS, NT and pituitary gland	FB, HB, NT, MHB and OB	5´from opo; H3K4me1 and H3K4ac27
H6:09601	FB	FB	5´from opo; H3K4me1 and H3K4ac27
H6:09617	FB, MHB and CNC	MB, lens, HA and low expression in FB	5´from opo; H3K4me1 and H3K4ac27
H6:09949	FB, DRGs and PF *primordium*	FB and PF	Within opo intronic region; H3K4me1 and H3K4ac27
H6:10137	**Presumptive NR**	**Dorsal NR**	**3' from opo; H3K4me1 and H3K4ac27**
H6:10377	DRGs, Pn, ears, and epidermis	DRG, ears and epidermis	3' from opo; H3K4me1 and H3K4ac27s
H6:10431	FB, BA, DRGs and ears	FB, BA, DRGs, ears, heart chamber	3' from opo; H3K4me1 and H3K4ac27

Table 3. *Opo locus* enhancers expression patterns. GFP expression patterns during zebrafish development, and their putative regulatory influences in tissue development: Central nervous system (CNS); Midbrain hindbrain boundary (MHB); Forebrain (FB); Branchial arches (BA); palate

(PL); Neural tube (NT); Hindbrain (HB); Olfactory bulbs (OB); Pectoral fins (PF); Dorsal root ganglia (DRGs); Neural retina (NR); Pronephros (Pn).

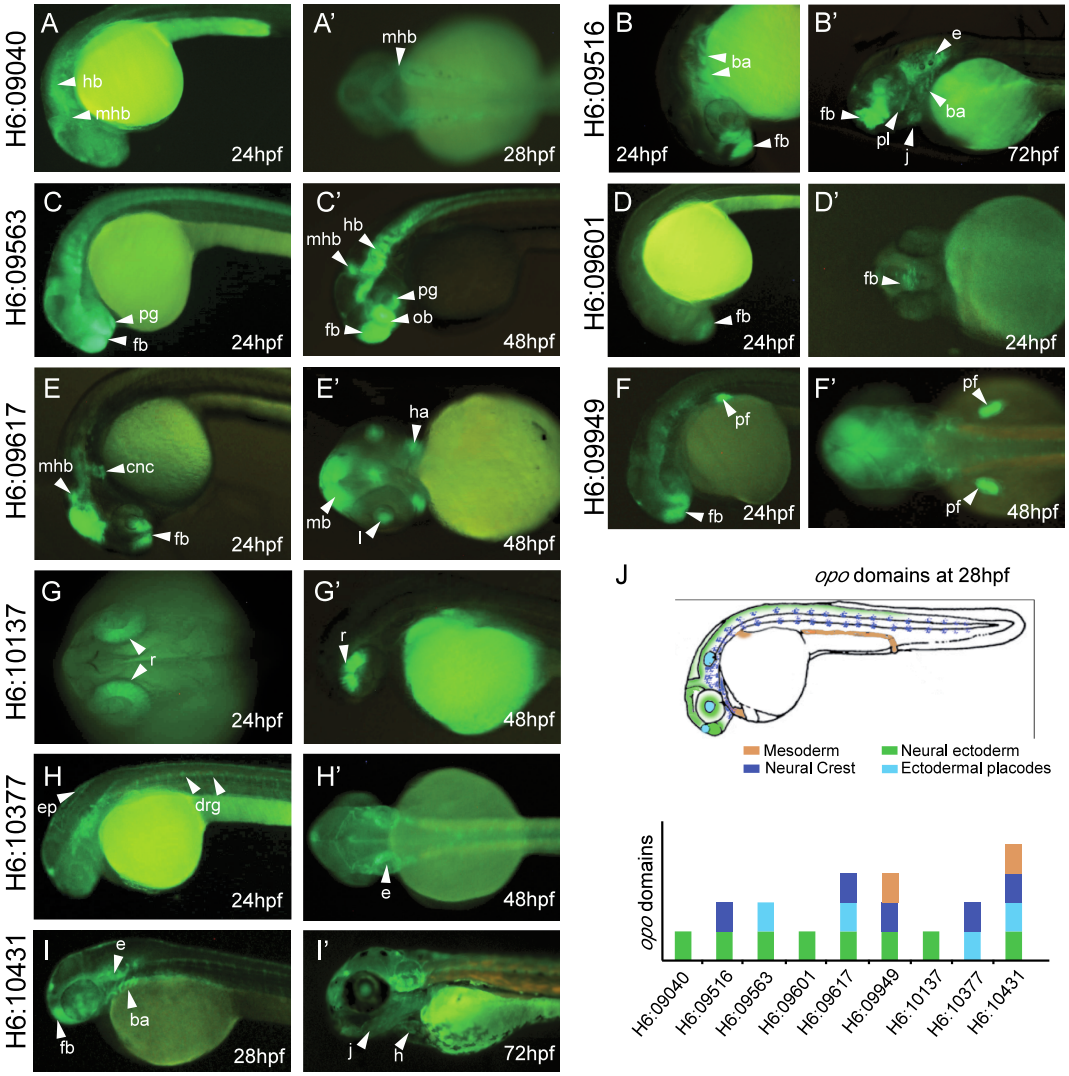
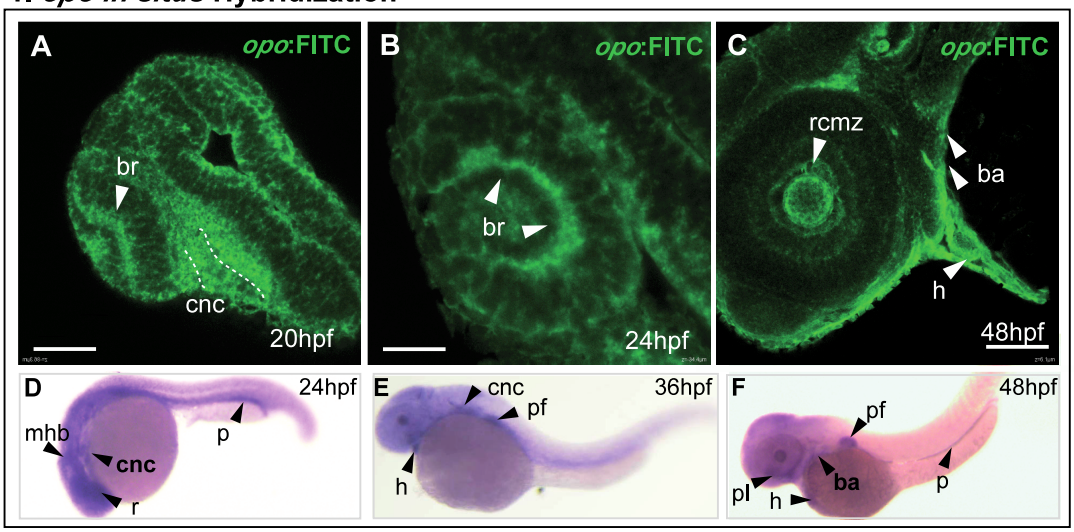


Figure 13. GFP expression pattern of *opo locus* enhancers. A-F) Arrowheads: midbrain-hindbrain boundary (mhb); branchial arches (ba); forebrain (fb); ear (e); palate (pl); jaw (j); hindbrain (hb); midbrain (mb); pituitary gland (pg); olfactory bulbs (ob); hyoid arch (ha); lens (l); pectoral fin (pf); retina (r); dorsal root ganglia (DRG); epidermis (ep); heart (h). J) The zebrafish model is colored according to the sum of the enhancers' GFP patterns. The total sum of the *opo* enhancers patterns appears to recapitulate the *opo* expression pattern as described in medaka (Martinez-Morales et al., 2009; Mertes et al., 2009) and mouse (Mertes et al., 2009).

The expression patterns of the individually isolated enhancers were compared to the *opo* expression pattern during zebrafish development, as revealed by *in situ*

hybridization. As an additional reference, a stable transgenic line harboring GFP under the control of the *opo* promoter region (*opo*5kbup:GFP) was generated. This 5 Kb genomic region was isolated from the compact medaka genome, taking advantage of the precise definition in medaka of the transcription starting site of the gene, as determined by RACE experiments in Martinez-Morales et. al. 2009. The expression patterns obtained from both analyses were similar (Figure 14 and Supp. Fig. 3) although not completely overlapping. Expression starts at around 16hpf in the optic vesicle and central nervous system. At 24hpf, *opo* is expressed in cranial neural crest derivatives, inner ear, pronephros, and particularly in the anterior CNS, including retina, forebrain and midbrain-hindbrain boundary. By 48hpf, *opo* is present in craniofacial structures derived from cranial neural crest; like jaw, palate, and branchial arches. The comparison of the GFP pattern of each enhancer with the global expression *of opo* confirms that, collectively, the enhancers recapitulate *opo* expression during development (Figure 14-J).

1. *opo in situs* Hybridization



2. *opo* promoter region

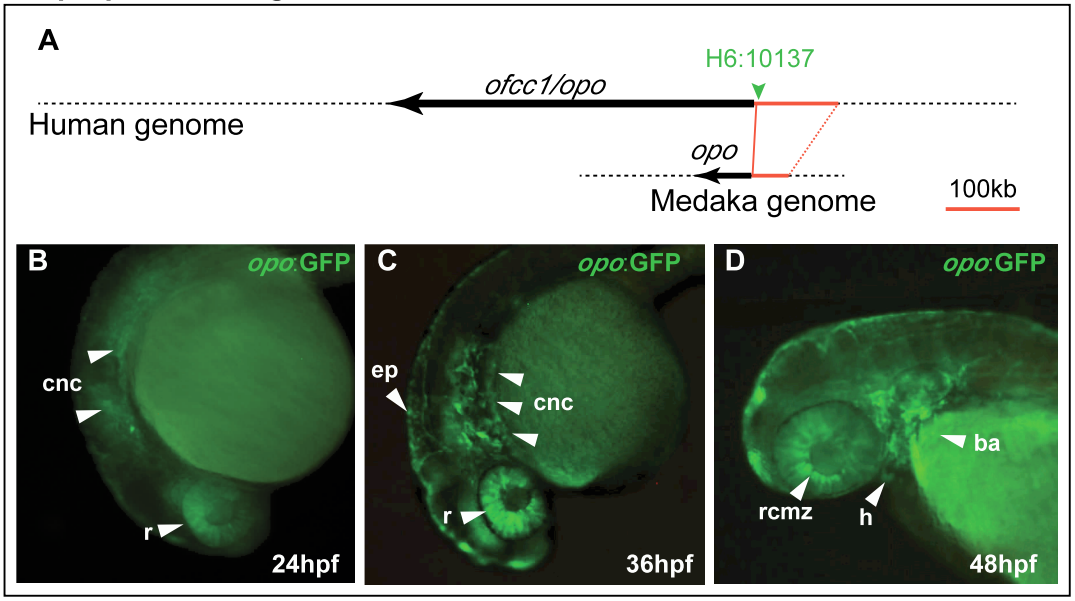


Figure 14. *Opo* expression pattern analyses. Panel 1: A-C) whole mount fluorescent *in situ* hybridizations (FISH); D-F) whole mount *in situ* hybridizations. Dashed lines: cranial neural crest. Arrowheads: (br) basal side of the presumptive neural retina; (cnc) cranial neural crest cells; (ba) branchial arches; (h) heart; (rcmz) retina at ciliar marginal zone; (mhb) midbrain-hindbrain boundary; (p) pronefros; (pf) pectoral fin; (pl) palate. Panel 2: A) *opo* promoter region amplified from the medaka genome contains the H6:10137 enhancer (green). B-D) GFP expression pattern of the *opo* promoter region stable line (Tg(*opo*5kbup:ZED). B) GFP expression in neural retina and in cranial neural crest at 24hpf; C) GFP expression in neural retina cranial neural crest and epidermis at 36hpf; D) GFP expression pattern at retina ciliar marginal zone, branchial arches and heart.

The 5 kb of opo promoter region includes H6:10137 enhancer inside, which is expressed exclusively in the neural retina. Accordingly, *opo* prmoter region displays specific GFP expression in neural retina. Although, this 5kb region also dirves the GFP expression to the cranial neural crest, but not in forebrain, midbrain, pronephros or fins. We could not find any cranial neural crest enhancer in this regions suggesting that we were unable to identify all regulatory elements using phylogenentic footprinting and epigenetic marks criteria.

The majority of the expression patterns associated to the enhancers were not exclusive in one tissue, but were instead expressed in multiple tissues at the same time. Also, multiple enhancers were expressed in the same domains, suggesting a combinatorial activity for the transcriptional control in specific tissues, such as in the forebrain and branchial arches (Figure 15).

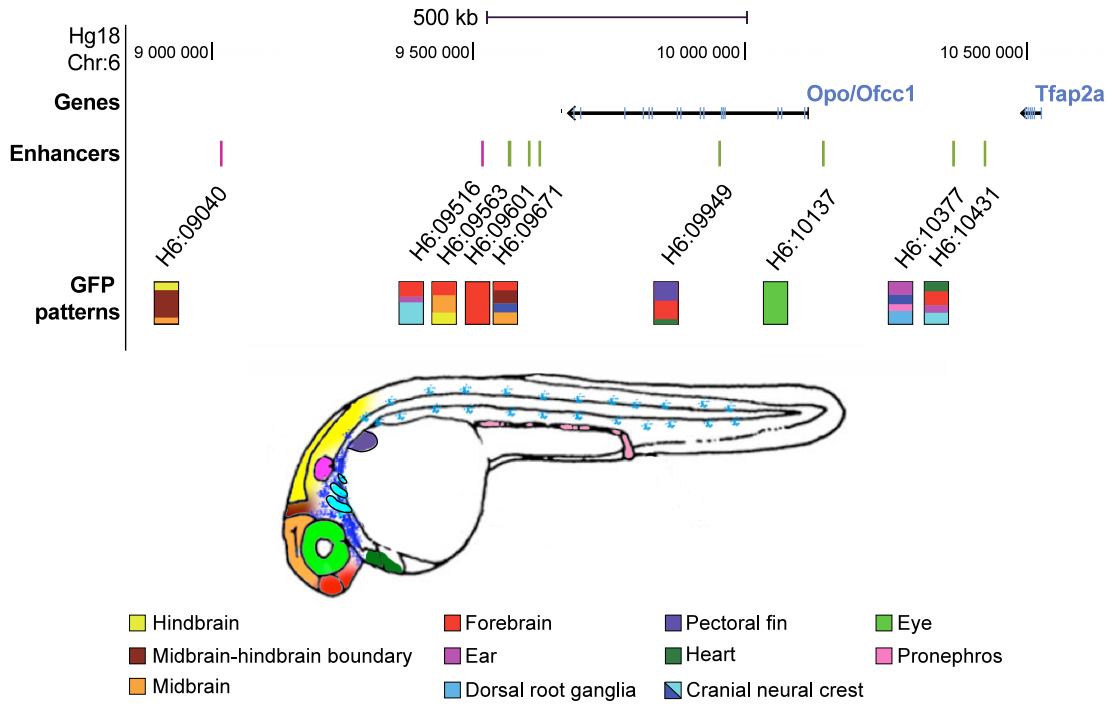


Figure 15. Distribution of the enhancers along the *opo* landscape. Representation of the enhancers GFP expression patterns in zebrafish at 28hpf. Multiple enhancers are expressed in the same embryo domains. Two enhancers (H6:09601 and H6:10137) are expressed in only one tissue.

In summary, the new enhancers found in the *opo locus* drive GFP expression primarily in the CNS and in cranial neural crest derivatives. Interestingly for this thesis, one of the CNS enhancers (H6:10137) was expressed exclusively during the neural retina development. This eye-specific enhancer, H6:10137, is contained within the analyzed *opo* promoter region and drives the expression of the GFP reporter in an remarkable pattern, which appears first at 16hpf, coinciding with the onset of optic cup folding, and remains "active" until 46hpf, when the eye is completely formed.

3. Nuclear environment of *opo* locus

If the enhancers identified in the *opo locus* are indeed regulating *opo* expression, they are expected to establish physical contact with the promoter of the gene. To explore the nuclear environment of *opo locus*, we employed the Circularized Chromosome Conformation Capture technique followed by deep sequencing of the obtained fragments (4c-seq). This technique allows the the genome-wide detection of the interactions between *opo* promoter (viewpoint) and its surroundings (Figure 16).

For the 4C analysis we use mouse genetic material at stage E9.5, corresponding to the onset of expression of *opo* during development (Mertes et al., 2009). Ideally we should use genetic material from zebrafish or human, taking in account that our study is done using human sequences and using zebrafsh as model organism. We did not use human genetic material because it is inaccessible. The zebrafish genentic material is available to get, however the annotation of the zebrafish genome is fragmented for the *opo locus*, what impairs the posterior analysis of the results.

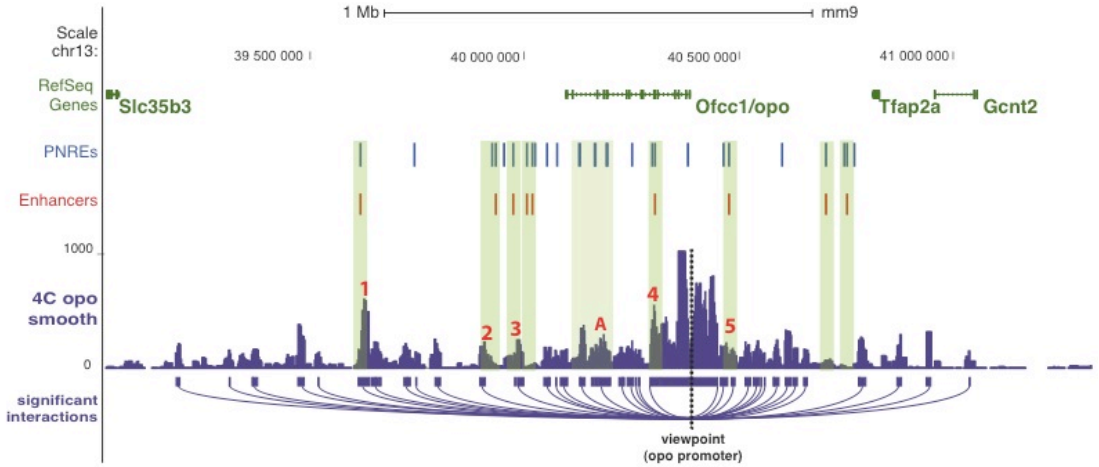


Figure 16. 4C analysis in *opo locus*. 4C analysis with a fixed fragment (dashed black line) in *opo* shows that opo promoter contacts (light green bars) with 5 of the 9 enhancers indentified in *opo locus* (red numbers: 1-H6:09040; 2- H6:09516; 3- H6:09563; 4- H6:09949; 5- H6:10137); Another elements inactive (H6:09815; H6:09892) and with no tissue specific enhancer activity (H6:09813) were also found to contact with *opo* promoter (red letter: A)

The analysis of the 4C data revealed that *opo* promoter establishes contacts with the majority of the previously identified *opo locus* enhancers. A large fraction of the contacts are mapped 3' from *opo* promoter and it was observed that 5 out of the 9 enhancers establish significant contacts with *opo* promoter. Additionally, the global interaction map of the *opo* promoter spands over a 1.5Mb genomic window covering the entire *opo* l*ocus* and containing the collection of identified PNREs .

These findings indicate that the majority of the newly identified *opo locus* enhancers maintain physical contacts with *opo* promoter, which suggests a regulatory influence of these elements over *opo* expression.

The 4C analyses were also preformed using the H6:10137 enhancer as viewpoint (Figure 17). The obtained data revealed that H610137 enhancer is contacting the *opo* promoter, two *opo* intronic regions and one enhancer (H6:09601) . The fact that H6:10137 enhancer contacts other enhancers at the same time that is contacting the *opo* promoter suggest a complex architecture of the *opo locus*.

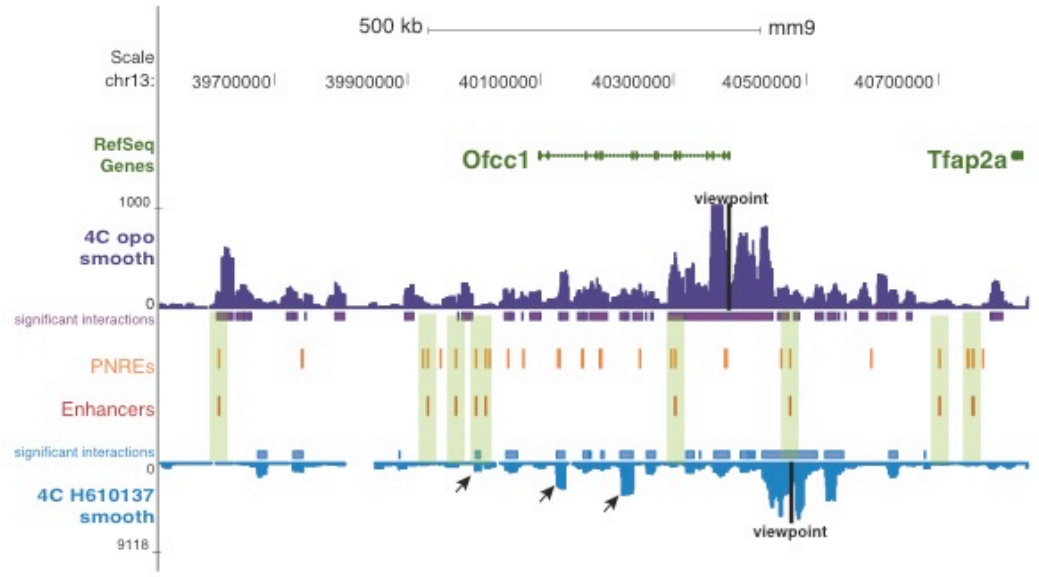


Figure 17. 4C analysis in *opo locus* landscape using *opo* promoter and H6:10137 as viewpoints. Fixed fragments used as viewpoints are represented as black lines. The interactions between the PNREs (orange) and the enhancers (red) with the opo promoter and H6:10137 enhancer are marked with light green bars. The arrows mark the interaction of H6:10137 with both intronic regions and H6:09601 enhancer

Given the known function of *opo* during eye development, the expression of H6:10137 enhancer suggests that it may be a key regulator of *opo* expression in the retina. Thus, in order to identify putative up-stream regulators of *opo*, the H6:10137 sequence was analyzed in detail for the presence of transcription factor (TF) DNA-binding sites.

4. Search for transcription factor binding sites in eye enhancer H6:10137

To search for conserved transcription factor (TF) DNA-binding sites within the H6:10137 sequence, we used MultiTF alignment tool from the ECR browser (Ovcharenko et al., 2004). Through this method, a list of 56 TFs binding sites conserved between human (hg18) and xenopus (*xenopus tropicalis*) was obtained (Supp. Tab. 5). Additionally, the JASPAR Database (Bryne et al., 2008) identified a list of 20 TFs binding sites with a score $\geq$ 10 (Supp. Tab. 6). The combined list of transcription factors was then examined in order to verify which TFs shared a common expression pattern with H6:10137. To this end the expression annotations described in The Zebrafish Model Organism Database (Sprague et al., 2003) were

used as a reference.

These investigations revealed a number of transcription factors expressed in the retinal tissue, such as Vsx2, Rx, Vsx1, Brca. Among them, the homeobox transcription factor *Vsx2* (also called *Chx10* in mammals) showed a neural retina specific expression pattern (Liu et al., 1994; Vitorino et al., 2009) that resembled most the spatio-temporal activation of H6:10137 (Figure 18-A,B). The comparison of the data obtained with ECR and JASPAR databases, indicated the presence of two conserved binding sites for Vsx2 (Zou and Levine, 2012) in H6:10137: BS1:142-151bps, detected both with ECR and JASPAR; and BS2: 678-687bps, detected only with ECR. These binding sites were then aligned for three representative species using *ClustalX*, (Figure 18- C, D). The resulting alignments further confirmed that the two binding sites for Vsx2 in the H6:10137 sequence are conserved from mammals to teleosts.

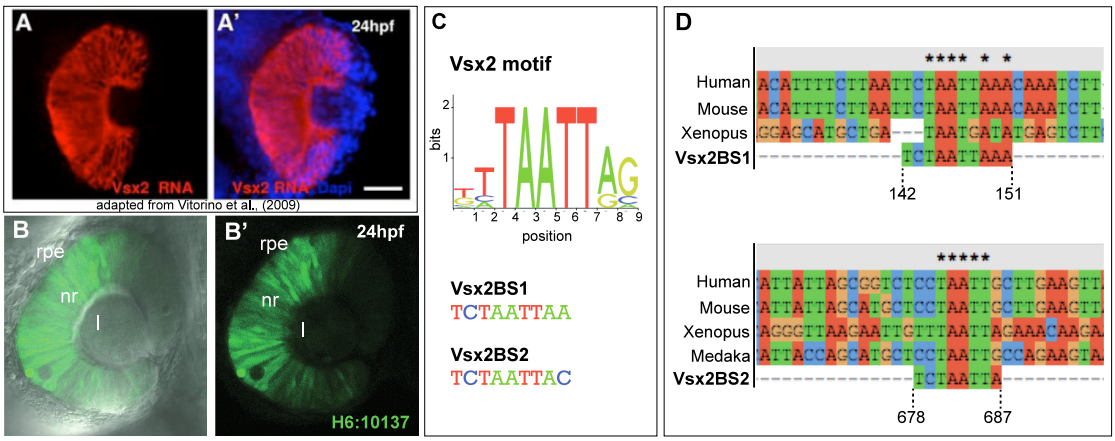


Figure 18. H6:10137 enhancer activity and Vsx2 binding sites. A) *Vsx2* RNA expression in the neural retina at 24hpf (adapted from Vitorino et. al., 2009); B) GFP expression pattern of H6:10137 in the neural retina at 24hpf: (l) lens; (nr) neural retina; (rpe) retina pigmented epithelium; C) Vsx2 motif (Zou and Levine, 2012) and Vsx2 DNA-binding sites within the H6:10137 sequence; D) Multispecies alignment of H6:10137 enhancers show sequence conservation of BS1 and BS2; Genomes versions: Human: Mar. 2006 (NCBI36/hg18); Mouse: July 2007 (NCBI37/mm9); Xenopus: Nov. 2009 (JGI 4.2/xenTro3); Medaka: Oct. 2005 (NIG/UT MEDAKA1/oryLat2).

The presence of these conserved binding sites indicate a possible interaction between the transcription factor Vsx2 and the *opo* eye enhancer, which is also consistent with the specific expression of Vsx2 and H6:10137 in the developing neural retina, but not in the adjacent RPE.

5. Vsx2 DNA-binding sites are necessary for H6:10137 enhancer activity

To investigate if Vsx2 binding sites BS1 and BS2 within the H6:10137 sequence are required for the enhancer activity of H6:10137, both sites were mutated to generate the mutated construct mutH6:10137 (Figure 19).

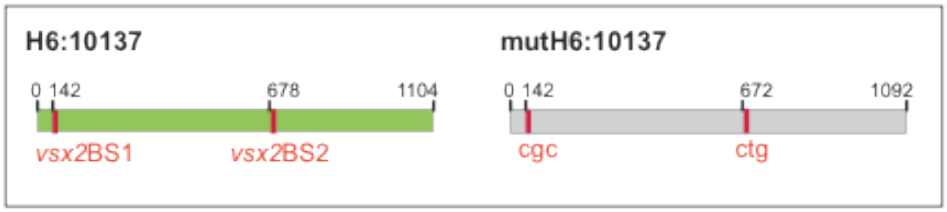


Figure 19. Mutagenesis of H6:10137. Generation of the mutH6:10137 construct by mutagenesis of Vsx2BS1 and Vsx2BS2 sites in the H6:10137 sequence.

The enhancer capacity of mutH6:10137 (activation of GFP expression in the neural retina) was compared with that of H6:10137 in mosaic embryos (F_0) produced in transient zebrafish transgenic assays. As positive control for transgenesis (RFP expression in somites and heart) was observed at 48hpf in both cases (Figure 20) and only positive embryos were scored. While 42% of the H6:10137 embryos positive for transgenesis showed typically elongated GFP positive neuroblasts in this tissue when analyzed at 24hpf, mutH610137 embryos do not present any GFP positive cells in the undifferentiated neural retina (Figure 20).

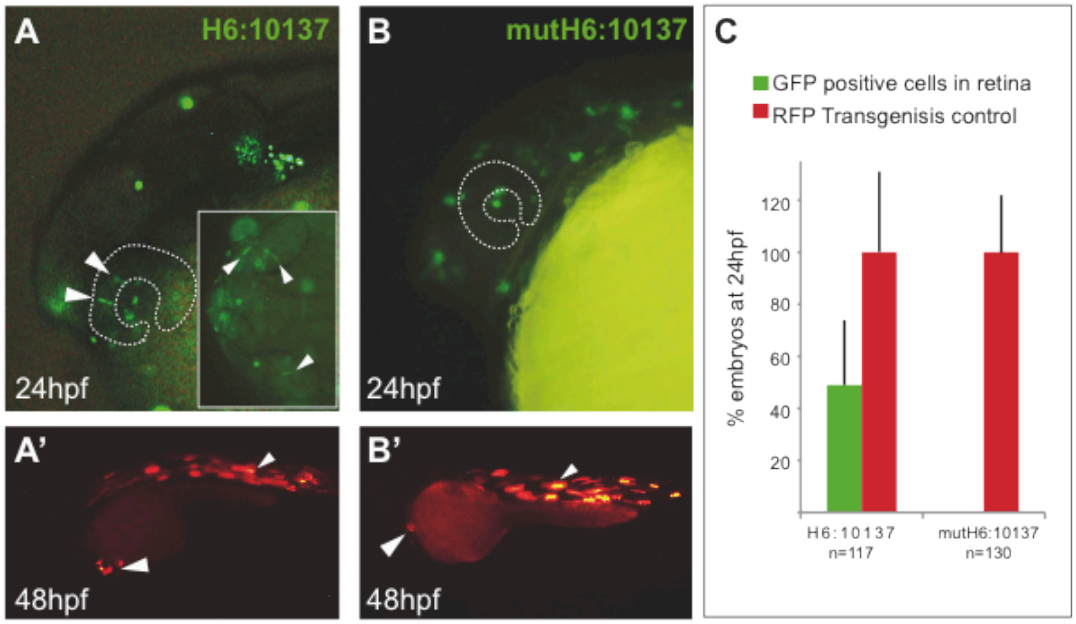


Figure 20. Transgenesis assays of mutH6:10137 and H6:10137. A-A') H6:10137 transgenic embryos (F_0 - mosaic) expressing GFP in two retina cells and RFP in heart and somites, (arrowheads); B-B') mutH6:10137 transgenic embryos (F_0 - mosaic) without GFP activity in retina cells and expressing RFP controls in heart and somites (arrowheads); C) 42% of the H6:10137 transgenic embryos had GFP positive cells in retina at 24hpf. Error bars indicate standard deviation of the mean.

Therefore, the mutation of Vsx2 BSs in the H6:10137 enhancer sequence abolished its enhancer activity in neural retina precursors. This suggests that the activity of the enhancer depends on Vsx2 binding.

6. Vsx2 regulates *opo* expression levels through its binding to H6:10137

We reasoned that if H6:10137 activity depends on Vsx2 binding, loss of *vsx2* function should also affect both H6:10137 activity and *opo* expression levels. To address this point, we designed splicing morpholinos against *vsx2*, to produce a knockdown condition.

Two splicing blocking morpholino oligonucleotides (MO) were designed: I1E2MO and E2I2MO, affecting the splicing of introns 1 and 2 respectively. Both MOs generate aberrant transcripts, as assessed by RT-PCR amplification from morphant embryos cDNA. Sequencing of these transcripts confirmed splicing disruption of Vsx2 mRNA in morphant embryos (Figure 21- A, B).

To evaluate the effects of Vsx2 knockdown on H6:10137 activity, both *vsx2*

morpholinos (*vsx2*MOs) were injected separately into one-cell stage into Tg[H6:10137] embryos (i.e. the stable H6:10137 transgenic line). GFP activity in Tg[H6:10137] morphant and control embryos was examined from 18 to 24 hpf. In addition to conspicuous morphological eye defects, which will be extensively discussed in the next section, *vsx2* morphants showed a dramatic decrease of the levels of GFP signal. In fact, GFP signal could only be detected in 20-30% of the injected embryos (Figure 21-D, G).

These results show that a reduction of functional *vsx2* in the neural retina impinges on the activity of the H6:10137 enhancer; thus indicating that Vsx2 is necessary for its regulation.

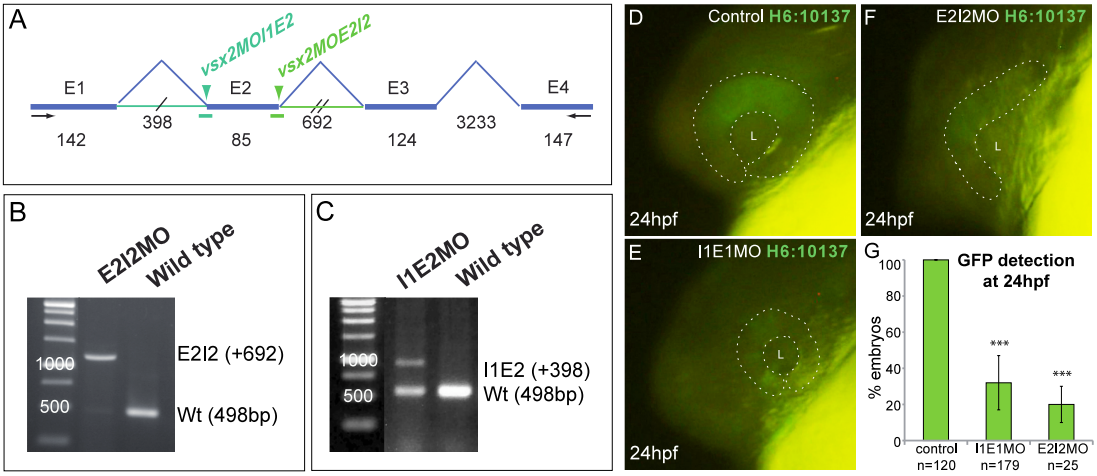


Figure 21. *Vsx2* knockdown using splicing morpholinos. A) The zebrafish *vsx2* transcript. Exon 1 (E1); Exon 2 (E2); Exon 3 (E3); Exon 4 (E4); Introns are represented as dashed lines; the length of both Introns and Exons is represented above. RT-PCR primers are represented as two arrows at transcript ends. Vsx2MOs are represented in green. B-C) RT-PCR of E2I2MO and I1E2MO injected embryos at 24hpf and respective controls (wild type embryos) showing the morphants' aberrant transcripts with the inclusion of Intron 2 in E2I2MO (+692bp), and Intron 1 in I1E2MO (+398bp). In both conditions the amount of native transcript (Wt: 498bp) is reduced when compared to the controls, however the reduction is more evident in the E2I2MO embryos. D-F) GFP activity in *vsx2* morphant embryos is reduced when compared to the controls (the eye is marked with a dashed line). G) GFP detection in morphant embryos is significantly reduced when compared to the controls, using a Student *t*-test (*p*-value <0,0001). Error bars indicate standard deviation of the mean.

Following the previous line of arguments, if Vsx2 is essential for H6:10137 enhancer activity during eye development, and H6:10137 is likely to regulate *opo* expression, it is expected that *opo* expression depend on *vsx2* levels.

To test this, we generated *vsx2* gain and loss of function conditions we either injected zebrafish *vsx2* RNA (0,1 – 0,3 ug/ul *per* embryo) or the above described morpholinos, respectively. Then we used RT-qPCRs to quantify *opo* levels in *vsx2* knockdown and *vsx2* over-expresing embryos at 24hpf. These quantifications showed a significant reduction of *opo* expression in *vsx2* knockdown embryos, when compared to the controls. This reduction was even more evident when only the heads of knockdown embryos were used as sample for the RT-qPCRs. As expected, the expression levels of *opo* were increased in *vsx2* over-expressing embryos with respect to the controls (Figure 22).

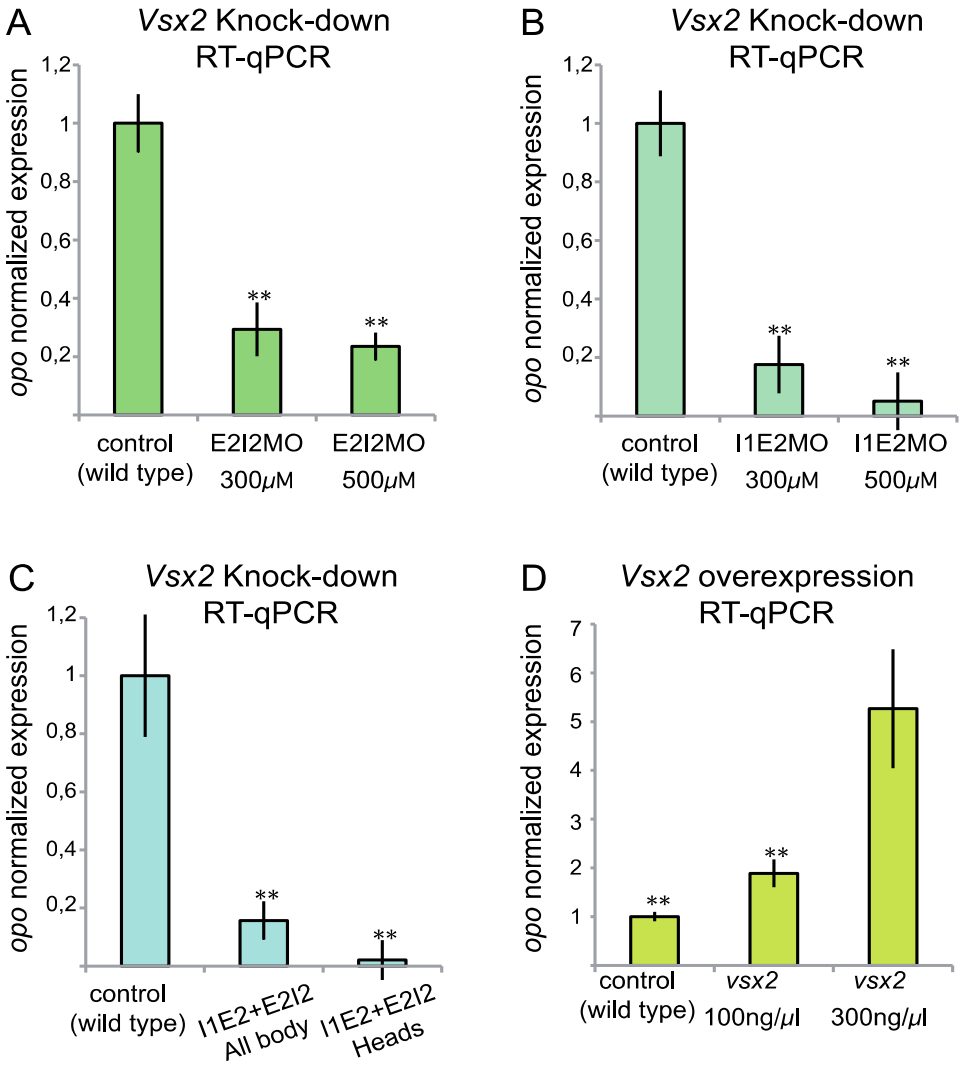


Figure 22. *Opo* expression levels quantifications by RT-qPCR. *Opo* expression levels are significantly reduced in *vsx2* morphant embryos in a dose-dependent manner (A-B). This reduction is more pronounced when only the heads of co-injected embryos with both *vsx2*MOs were used for the quantification (C). *opo* expression levels in embryos over-expressing *vsx2* are significantly higher when compared to controls (D). N=3; Student *t*-tests. Error bars indicate standard error of the mean (SEM).

According to the previous results, Vsx2 regulates the activity of the H6:10137 enhancer and consequently *opo* expression levels. To confirm that this regulation is direct we performed ChIP-qPCR assays (Chromatin Immunoprecipitation followed by quantitative PCRs). To this end we generated a *GFP_Zfvsx2* fusion and its *corresponding* RNA was injected into one-cell stage embryos (see methods). Then ChIP assays were performed using the heads of 300 injected embryos and a ChIP grade antibody against GFP. In qPCRs experiments, primers for the housekeeping *ef1a* gene were used as a negative control, once confirmed that the amplified sequence had no Vsx2 binding sites. As positive control we used primers for the *vsx1* promoter, previously described as a Vsx2 target in Clark et al., (2008). These assays confirmed *in vivo* the binding between Vsx2 and the enhancer H6:10137 (Figure 23).

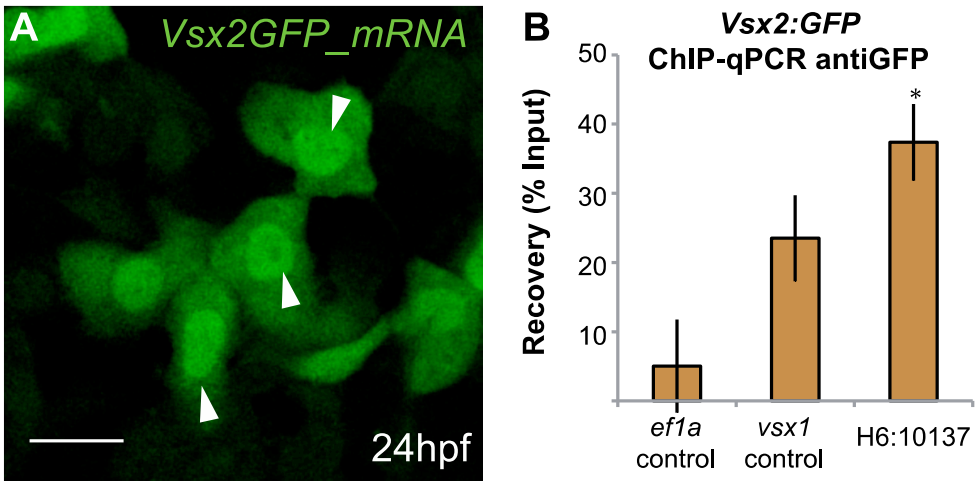


Figure 23. H6:10137 is a direct target of Vsx2. A) Confocal microscopy image of epidermal cells injected with *vsx2*GFP-mRNA, showing the nuclear localization of the mRNA (arrow heads) at 24hpf. B) Using chromatin extracts of *vsx2GFP*-mRNA-injected embryo heads a ChIP-qPCR was performed, using anti-GFP. The recovery of the H6:10137 sequence was significantly increased when compared to the negative control (*ef1a*) (n=2). One Way ANOVA; error bars indicate SEM.

These findings confirm that the H6:10137 enhancer is a direct target of the transcription factor Vsx2 and hence indicate that *vsx2* is an upstream positive regulator of *opo*.

7. *Vsx2* disruption causes optic cup folding defects

To investigate the role of *vsx2* during eye morphogenesis, optic cup folding was analyzed in *vsx2* morphants. Embryos were obtained from several independent injections of the above mentioned *vsx2*MOs (I1E2MO and I2E2MO). The *p53* morpholino (*p53*MO) was co-injected with *vsx2*MOs at 100uM to avoid the reported increased apoptosis artifacts (Langheinrich et al., 2002). Control embryos were either injected with 100uM of *p53*MO or uninjected wild type embryos were used as a reference. Injections were performed on the background of the Tg*(vsx2.2*:GFPcaax) line, which was generated in our laboratory using the Medaka *vsx2.2* promoter and GFPcaax as reporter gene. This line, in which neural retina membranes are fluorescently labeled, facilitates the analysis of morphant phenotypes. As a transgenesis control, this line also expresses GFP in the heart (Kwan et al., 2007).

The phenotype analysis shows that both *vsx2*MOs were able to affect eye development when compared to the controls, producing similar eye phenotypic defects. Morphant embryos display two main morphological defects: a reduction of eye size (microftalmia) and the malformation of the optic cup, now displaying large ventral openings. 23% I1E2MO morphant embryos did not show any eye defects, 66% of the morphants showed eye morphological defects and 12% of anoftalmic (n=46, *p*-value <0,0001). For the E2I2MO morphant embryos 28% did not show eye defects, 57% displayed eye morphological defects and 15% were anoftalmic (n=50, *p*-value<0,0001). In the controls it was not observed morphological eye defects (n=40).

Micorftalmia and ventral opening phenotypes overlap in 40%-50% of the affected embryos. The complete absence of the eye (anoftalmia) was observed in the most severe cases (see Figure 23). Interestingly, the E2I2MO morphant embryos showed stronger eye defects than the I1E2MO-injected embryos (see Figure 24- E, F). This correlates with our previous finding showing that in fact E2I2MO induces a deeper knockdown of vsx2 mRNA than I1E2MO.

To test for the specificity of the morphant phenotype, *vsx2*-mRNA was coinjected

with E2I2MO. The morphant phenotype was significantly rescued in 47% of the co-injected embryos (73% No eye defects, 27% Eye morphological defects; N=53) when compared to embryos injected only with E2I2MO (28% of No eye defects, 57% of Eye morphological defects, 15% of anoftalmic, n=50) (Figure 24- E, Supp. Tab. 7).

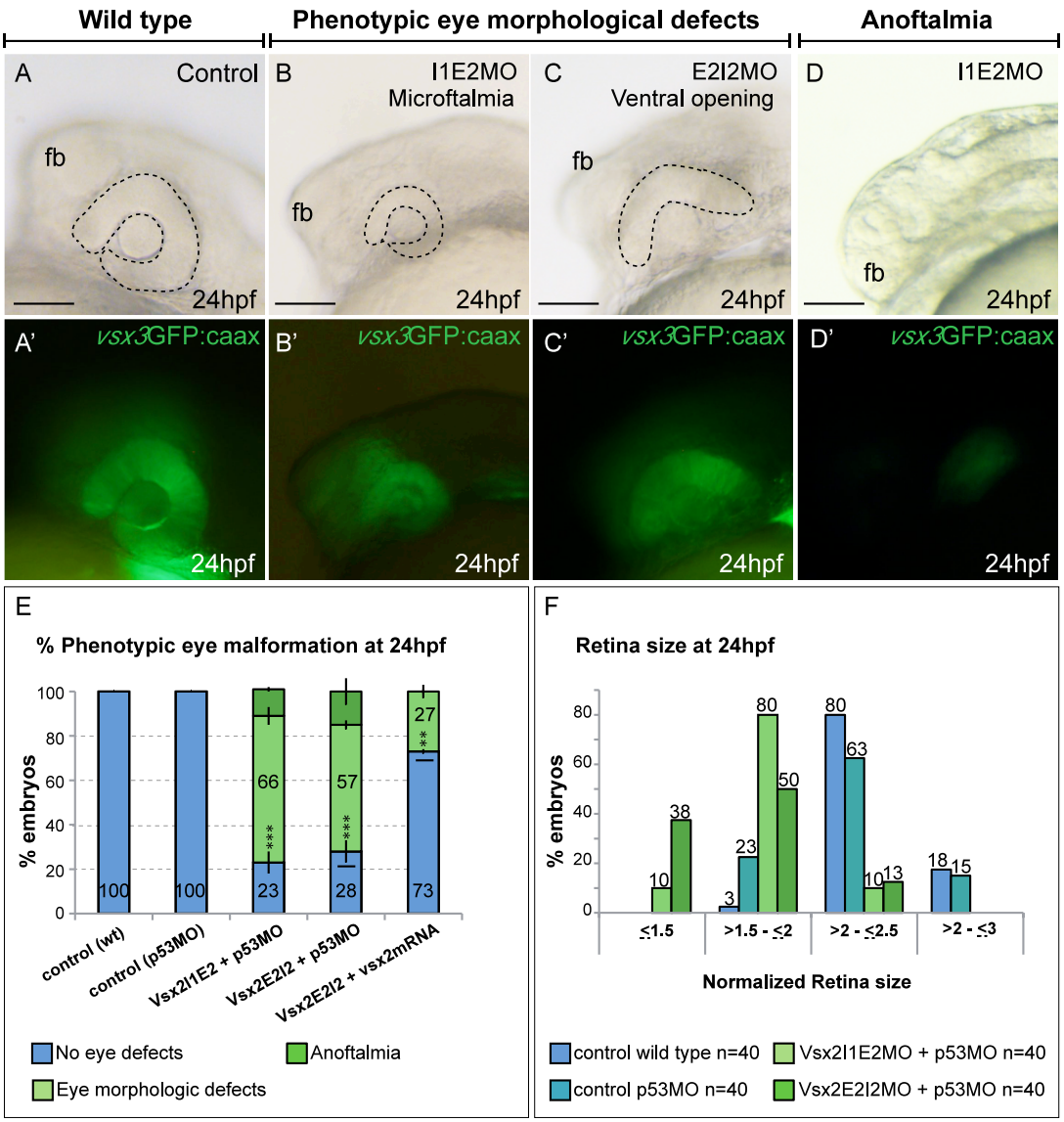


Figure 24. Phenotypical defects of *vsx2*MOs-injected embryos. A-A') wild type control embryos showing no eye defects; B-C) Eye defects in *vsx2*MO-injected embryos (retinas are marked with a dashed line), miroftalmia (B) and ventral opening (C); D) *vsx2*MO-injected embryos with anoftalmia; E) *vsx2* morphant embryos show significant phenotypic eye malformations when compared with controls. In the co-injected embryos with E2I2MO and *vsx2*-mRNA, the % of embryos showing ocular malformations is significantly reduced when compared to embryos injected with E2I2MO; F) In *vsx2* morphants the retina is reduced when compared to the controls.

Error bars indicate standard deviation.

The finding that *vsx2* loss can result in microftalmia and anoftalmia had already been observed in previous studies in zebrafish at later stages (Vitorino et al., 2009) and also in *vsx2* Knockout mice models (Burmeister et al., 1996). Although our analysis was performed at earlier stages (20 to 25hpf), when such malformations are harder to identify, we were also able to observe eye reduction defects in over 30% of knockdown embryos (Figure 24, Supp. Tab. 8). Interestingly, retinae in morphant embryos were still ventrally opened at 24hpf, when the folding of the optic cup should be completed (Figure 24-C).

To analyze this phenomenon in a quantitative manner, the angles of ventral openings were measured in over 100 embryos *per* morpholino, at 24hpf. This analysis revealed that around 70% of *vsx2* knockdown embryos displayed ventral openings in the retina, and these openings ranged in angle between 21° to 120°. This is much in contrast to control embryos that showed ventral openings in a range of 0° to 20°. E2I2MO injections produced even stronger phenotypes, with 23% of morphant embryos displaying ventral openings wider than 80° when compared to I1E2MO morphant embryos (12% of ventral openings wider than 80°) (Figure 25, Supp. Tab. 9).

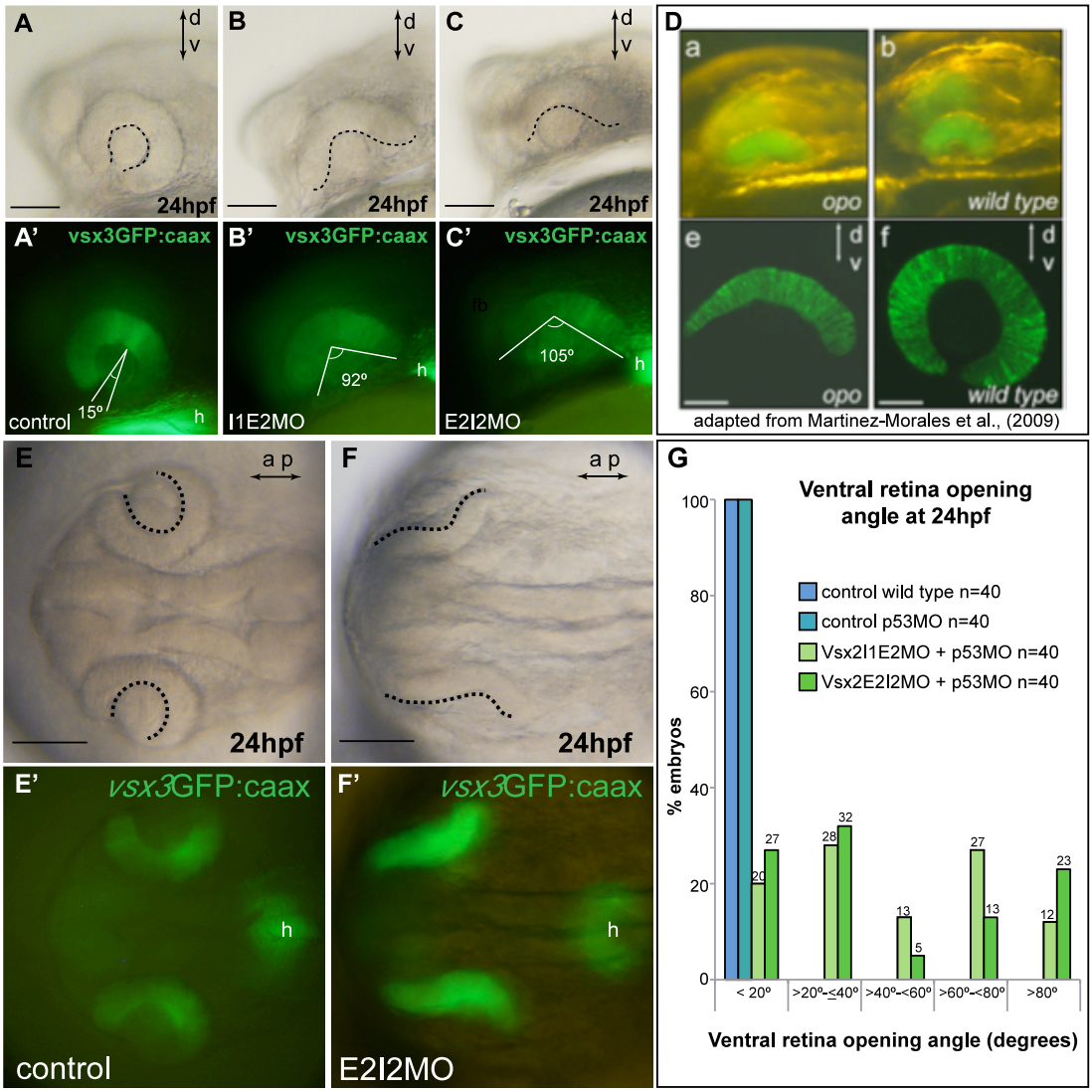


Figure 25. *Vsx2* morphants: Ventral opening retinas measurements. A-A') control wild type embryos; heart (h); B-C') *vsx2*MO morphant embryos showing ventral opening retinas higher than 80°; D) Medaka *opo* mutant expressing *rx2*::mYFP (a and e) wild type *rx2*::mYFP medaka embryo (b and f) (Martinez-Morales et al., 2009); E-E') control wild type embryos (dorsal view). F-F') E2I2MO-injected embryo showing retinae ventral opening (dorsal view). The basal curvature of the retinae is marked with a dashed line. G) Measurements of the ventral opening angles in the *vsx2*MOs morphant embryos and controls at 24hpf.

We can therefore conclude that *vsx2* is necessary for the folding of the optic cup, being the morphological defects produced by its knockdown in the most severe cases similar to the defects observed in *opo* mutant eyes.

Together these results indicate that *vsx2* is an upstream regulator of *opo* via the H6:10137 enhancer. Defects in *vsx2* function result in a down-regulation of *opo* expression, and consequently its function in the folding of the optic cup is compromised. Previous

analyses did not establish a clear connection between the patterning of the optic cup and its morphogenesis. The results here presented show that *vsx2*, by controlling the expression of *opo*, provides an important link between both phenomena.

IV. Discussion

This work provides new insight into the regulation of the *opo* gene in vertebrates. Here, we identify and describe a battery of enhancers that regulate *opo* function during craniofacial and eye development. Additionally, we show that during optic cup folding in zebafish, Vsx2 regulates *opo* via the H6:1037 enhancer.

1. The *Opo* regulatory landscape

Organogenesis requires the dynamic control of gene expression, which in turn involves the activity of non-coding *cis*-regulatory elements (CREs) spread over large genomic distances. The regulatory landscape is defined by the interactions established by the CREs and specific gene promoters. Understanding the logic of regulatory landscapes is essential to interpret how basic genetic programs operate in the context of both development and disease.

In this work we focus on *ojoplano*, a developmental gene essential for both eye morphogenesis and craniofacial development. We dissect out the CREs that regulate its expression during these key developmental events; we make correlations between these new regulatory elements and the control of *opo* expression, and finally we suggest potential links for these new elements with human hereditary diseases.

Our comprehensive expression analysis by transgenesis of a representative collection of human putative non-coding cis-regulatory elements (PNREs) in zebrafish embryos, revealed the presence of nine tissue specific enhancers expressed in the *opo* domain. The majority (six) of these *opo* enhancers were expressed during the development of craniofacial structures such as the jaw and the palate. They were also expressed in the central nervous system, including the eye, and in sensory organs, like the olfactory bulbs and dorsal root ganglia. When the expression of *opo* (Martinez-Morales et al., 2009; Mertes et al., 2009) and its enhancers were analyzed, we observed that when summed together the enhancers expression fully recapitulates the *opo* expression pattern. This indicates that this short collection of enhancers might account for a substantial part of the *opo* gene regulation.

We also analyze the expression of a 5kb region upstream of the *opo* start site (opo_5kbup). This region, that includes the *opo* promoter and the H6:10137 enhancer, was also able to drive GFP expression to the neural retina, thus confirming the

regulatory influence of H6:10137 during eye development. *opo*5kbup:GFP was not able to recapitulate the full expression of the gene (e.g. there was no expression in the CNS, fins or sensory organs), however it was also expressed in the cranial neural crest. This result suggests the presence in this genomic region of cryptic neural crest enhancers, which our combined phylogenetic footprinting and/or epigenetic marks analyses failed to identify. The presence of additional enhancers in this region suggests that the regulatory complexity of the *opo locus* is even higher than anticipated. In addition, it also suggests the presence of shadow enhancers (Hong et al., 2008) that, together with the neural crest enhancers identified in this study, might drive redundant regulatory information.

Once we had defined the enhancers collection in the *opo* landscape, we performed a global analysis of the *opo* promoter interaction map by 4C (Circular Chromosome Conformation Capture). This analysis allows both the examination of the chromosomal architecture of an entire locus in a systematic manner, and also to validate specifically whether the identified enhancers are contacting the *opo* promoter or not. The data we obtained from this validation suggests that the majority of enhancers are establishing physical contacts with the *opo* promoter. Additionally, we found three elements (PNREs) not defined as enhancers in our analysis that also established significant contacts with the promoter: H6:09813, H6:09815; H6:09892 (Figure 13-A). In terms of enhancer activity, at least one of them, H6:09813, drove GFP activity. However, the GFP expression was not tissue-specific in the progenies of the different identified founders (Tab. 2). In these cases, it is possible that the sequences selected for transgenesis were not sufficient to induce tissue-specific enhancer activity, or, alternatively, these elements may need to be in their specific genomic environment to be able to drive tissue-specific activity.

Given the key role for *opo* during eye development that was recently described in our laboratory, we also decided to explore in detail the interaction map of the H6:10137 enhancer; which is expressed exclusively in the neural retina during eye morphogenesis. In agreement with the expression pattern analysis, the resulting data showed significant interactions between the H6:10137 enhancer and the *opo* promoter, suggesting that H6:10137 does directly regulate *opo* expression during eye development. More functional studies need to be performed to clarify if this enhancer is the only region required to control *opo* expression during eye

development. New genome editing tools (Jao et al., 2013; Xiao et al., 2013) open now the possibility to generate chromosomal lesions (insertions/deletions) in the enhancer H6:10137 sequence, and thus observe any resulting effects in *opo* expression and optic cup folding.

2. Opo enhancers and human hereditary diseases

The work presented here may also provide useful information about *opo* regulation in the context of human hereditary diseases. As mentioned before, in humans, *opo* is called *Ofcc1*, which stands for Oro Facial Cleft Candidate 1. The gene was given this name due to its association with orofacial cleft (OFC) syndrome (Davies et al., 2004; Murray and Schutte, 2004), a birth defect caused by abnormal facial development. The fact that *opo* mutant fish show defects in craniofacial structures like jaw, branchial arches and palate (Martinez-Morales et al., 2009), further strengthens the link between *opo* and this birth defect. Previous work form Davies et al., (2004) described three OFC cases in which the patients had independent translocations included within the *opo* locus, which were localized using previously published data and chromosome walking experiments. Importantly, some of these translocations do not compromise the coding region of the gene. They then generated representative clones for each identified breakpoint region: Breakpoint case 1: AL031906 and AL159986; Breakpoint case 2: AL031904; as well as Breakpoint case 3: AL031122 (Davies et al., 2004). With this information available, we have now investigated if any of the *opo locus* enhancers found in this study were localized within the identified translocation breakpoint cases. We established that enhancers H6:09516, H6:09563, H6:09601 and H6:09617 were all localized within the region that includes the breakpoint for case 2, and enhancer H6:10137 within that of the breakpoint for case 3 (Figure 26).

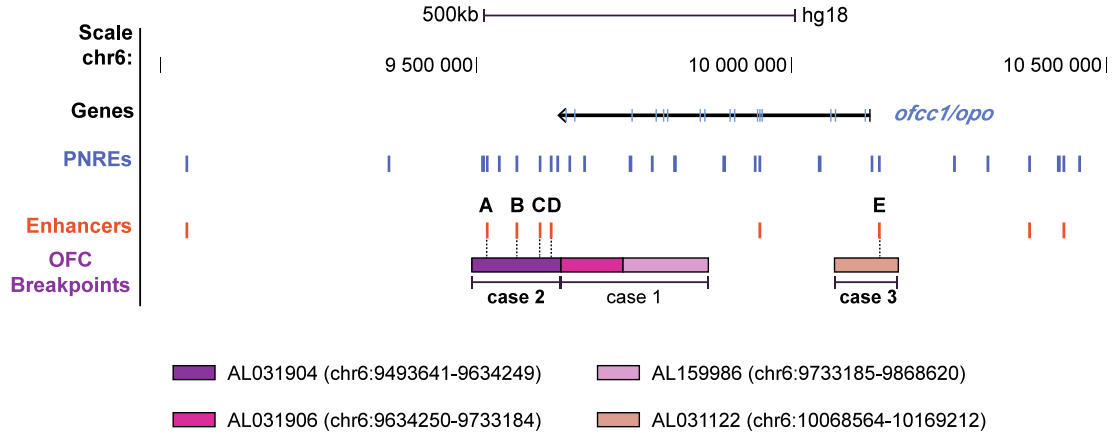


Figure 26. OFC breakpoints in the *opo* regulatory *locus*. Graphic representation of the *opo locus* (hg18 - chr6:8,961,047-10,550,262), showing the localization of the PNREs collection (blue), the newly identified enhancers (red) and the translocation breakpoints reported by Davies et al., in 2004 (pink shades). Enhancers H6:09516 (A), H6:09563 (B), H6:09601 (C) and H6:09617 (D) are localized within the region including the breakpoint corresponding to the OFC patient-case 2 (t(6;9)(p24;q22.3)). Enhancer H6:10137 (E) is localized within the region of the breakpoint corresponding to the OFC patient-case 3 (t(6;9)(p24.3;p23)) (adapted from UCSC Davies et al., (2004) and genome browser http://genome.ucsc.edu). Below, the accession number for each breakpoint-clone and its localization in the human genome is described.

The enhancers H6:09516 and H6:09617 are both expressed in facial derivatives of the neural crest; H6:09563 drives expression in the central nervous system and olfactory bulbs; while H6:09601 and H6:10137 are expressed exclusively in forebrain and neural retina respectively. Cranial neural crest (CNC) derivatives like branchial arches and jaw, as well as the central nervous system and eyes are all essential structures for the correct development of the craniofacial region (Santagati and Rijli, 2003; Kish et al., 2011). The fact that the enhancers H6:09516 and H6:09617 display expression in these structures suggests that they may contribute to craniofacial development. If this is the case, then the translocations between chromosome 6p24 and chromosome 9 would disrupt the cis-regulatory activity of the neural crest *opo* enhancers. Whether this is the causative mechanism explaining why the translocations associate to the craniofacial defects in patients with OFC needs to be formally determined.

As well as with OFC, the *opo locus* has also been associated with mental disorders, such as schizophrenia and Tourette's syndrome (TS) (Straub et al., 2002; Sundaram et al., 2011). There are several genetic markers located within *opo locus* associated with schizophrenia (D6S470; D6S940; D6S340), especially in family cases where the disease is related to cognitive defects (Straub et al., 2002; Hallmayer et al., 2005). Similarly, a novel single nucleotide variant was identified within the *opo locus* in two TS patients with different pedigrees (Comings et al., 1984; Sundaram et al., 2011). Interestingly, many of these genetic landmarks are located in non-coding regions, which suggests that causative mutations for these diseases could occur in cis-regulatory elements spread over the *opo locus*. Here, we observed that the majority of the 3' *opo* enhancers drive a strong GFP expression in the forebrain (H6:09516, H6:09563, H6:09617, H6:09601, H6:09949), suggesting a role for these enhancers in central nervous system development, and opening the possibility that a disruption of their activity could contribute to the development of the mental hereditary diseases. However, more functional and genomic studies need to be performed to investigate in detail the association between the *opo locus* and the development of schizophrenia or Tourette's syndrome.

3. The transcription factor Vsx2 binds to the enhancer H6:10137 to regulate *opo*

Vsx2 is a major gene in neural retina specification (Rowan et al., 2004). This TF is required to initiate the neural retina specification program by suppressing *mitf*, a retina pigmented-epithelium patterning gene, in the neural retina progenitor cells. The *vsx2*-mediated repression of *mitf* expression prevents *mitf* from implementing an RPE differentiation program in the developing neural retina (Horsford et al., 2005).
The mutation of *vsx2* is associated with microftalmia, which has been reported in humans and mice (ocular retardation mutant) and is caused by a decrease in the cell proliferation rate (Burmeister et al., 1996). Despite these observations, no morphogenetic defects have been reported in mammals so far. Previous knockdown studies in zebrafish at 48hpf also reported microftalmia and anoftalmia phenotypes (Barabino et al., 1997; Vitorino et al., 2009). One of these reports (Barabino et al., 1997), tangentially described early (24hpf) morphological defects in the zebrafish retina as a "flattened and triangular eye shape".

In this work, the analysis of the H6:10137 enhancer sequence revealed the presence of two conserved Vsx2 DNA-binding sites. Here, we tested the putative connection between *vsx2* and H6:10137 by mutagenesis assays for H6:101137, as well as with *vsx2* knockdown experiments. Furthermore, we confirmed the association between Vsx2 and the enhancer H6:10137 through the detection of physical interactions by chromatin immunoprecipitation (ChIP)-qPCR. Taken together, these experiments confirmed that H6:10137 enhancer activity depends on Vsx2 binding, and also that H6:10137 is a direct target of Vsx2 during optic cup folding.

In our *vsx2* morphant embryos, microftalmia, anoftalmia and retinal flat morphology were also present at early stages of eye development. This, in addition to previous experiments showing that the expression levels of *opo* were reduced in *vsx2* morphants, further support the hypothesis that *vsx2* plays a key role during optic cup folding, as an upstream regulator of *opo*. The similarities in morphology observed between ventrally-opened *vsx2* morphant and *opo* mutant retinae are also in line with this possibility. The evaluation of *opo* expression levels and eye morphology upon *vsx2* targeted mutation in a teleost model should provide an ultimate demonstration for this hypothesis in the future.

During early eye development, the neural retina precursors activate a genetic program that specify their cellular identity and is responsible for their morphological transformation during optic cup folding. It has been shown that this program acts in a tissue-autonomous manner, being to a large extent independent of neighboring tissues (Eiraku et al., 2011). Previous studies in medaka also demonstrate that optic cup folding occurs through a basal constriction mechanism, which is dependent of *opo* expression (Martinez-Morales et al., 2009). Clonal analysis shows that this folding mechanism depends on the basal constriction of the neural retina precursors, but not on the morphological transformation of the retinal pigmented epithelium precursors that passively fold towards their apical side (Martinez-Morales et al., 2009). Here, we suggest that *vsx2* may not be necessary only to confer identity to the neural retina precursors, but also to specify their morphogenetic behavior. It is likely that this cell identity program also comprises a 'morphogenetic identity' by the activation of effector genes like *opo*. In this sense, *vsx2* would provide the first direct link between

neural retina specification and morphogenesis at the onset of optic cup invagination in teleosts.

In contrast to our observations in zebrafish, eye morphological defects are not observed in *vsx2* mutant mice (Burmeister et al., 1996). Several explanations may account for this discrepancy. It is known that ocular retardation mutants can show a range of variability in the severity of the phenotype, predicting the presence of genetic interactors (Rowan et al., 2004). This observation suggests a complex network of interactions during retinal specification. In line with this, we observe that the overexpression of *vsx2* is not sufficient to induce ectopic expression of the H6:10137 enhancer. This finding may be explained by the fact that the eye enhancer also shows putative binding sites for other well known eye specifiers (i.e. *rx* and *crx*; Supp. Tables 5 and 6). These transcription factors may have different regulatory weight on H6:10137 in teleosts and tetrapods that could explain the phenotypic discrepancy.

V. Conclusions

1. The analysis of PNREs localized within the *opo* regulatory landscape revealed the presence of nine tissue-specific enhancers.

2. The identified enhancers were distributed equally at the 5´and 3´ sides of the gene; all were highly conserved and the majority (78%) show both H3K4me1 and H3K27ac marks in human cell lines and zebrafish.

3. The total sum of the enhancers' expression patterns recapitulates the expression pattern of the *opo* gene.

4. Only one enhancer was found to be expressed exclusively in early retina (H6:10137); most expression patterns were detected in the central nervous system, neural crest and craniofacial structures.

5. The study of the interaction profile of the *opo* promoter with its genomic neighborhood confirms that most of the identified enhancers contact physically the promoter.

6. The H6:10137 enhancer has two conserved binding sites for *vsx2*, and its activity depends on *vsx2*.

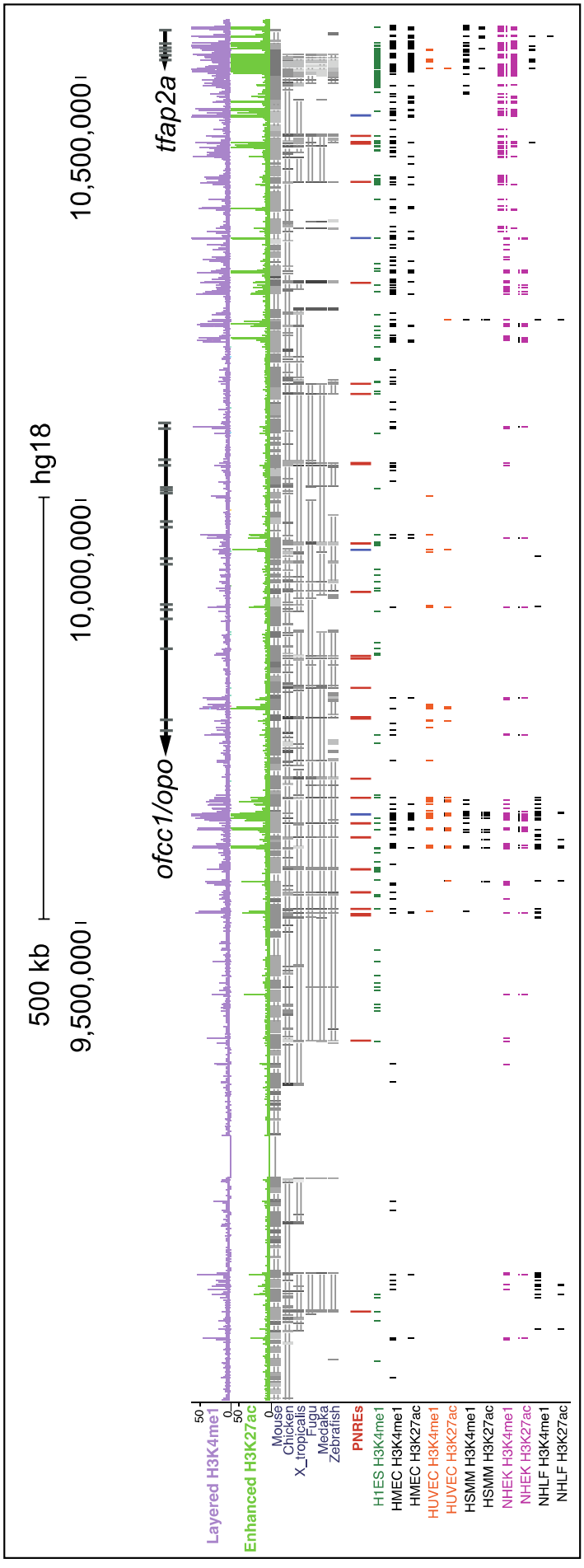
7. Vsx2 interacts physically with the enhancer H6:10137 *in vivo*.

8. *opo* expression itself depends on *vsx2*: it is up-regulated in *vsx2* over-expressing embryos, and down-regulated in *vsx2* knock-down embryos at 24hpf.

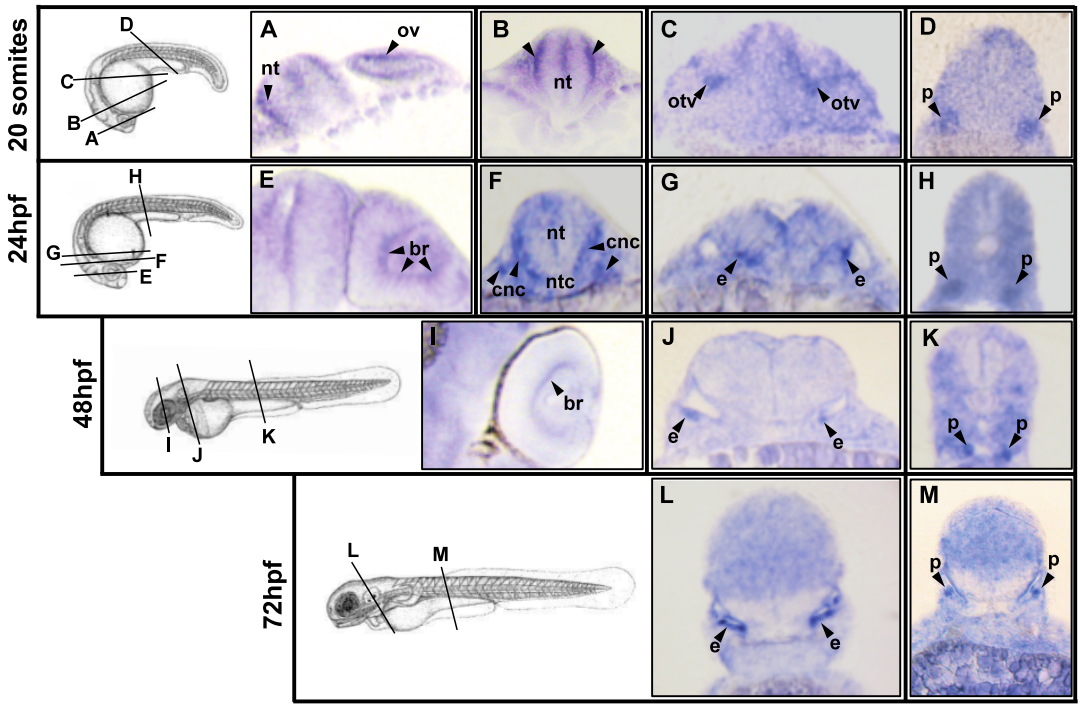
9. Knockdown of *vsx2* in zebrafish embryos results in a range of eye defects, including: ventral retina openings, microftalmia, and anoftalmia.

10. The optic cup morphogenetic defects observed upon *vsx2* morpholino interference resemble, in the most severe cases, the ocular phenotype in the medaka *opo* mutants.

IIV. Supplementary Material


500 kb
hg18
9,500,000
10,000,000
10,500,000
ofcc1/opo
tfap2a
Layered H3K4me1
Enhanced H3K27ac
Mouse
Chicken
X_tropicalis
Fugu
Medaka
Zebrafish
PNREs
H1ES H3K4me1
HMEC H3K4me1
HMEC H3K27ac
HUVEC H3K4me1
HUVEC H3K27ac
HSMM H3K4me1
HSMM H3K27ac
NHEK H3K4me1
NHEK H3K27ac
NHLF H3K4me1
NHLF H3K27ac

Supp. Fig. 1. The *opo* regulatory landscape. Localization: (hg18 2006) in chr6:8936346-10570697. Genes: *ofcc1* and transcription factor *AP2 alpha*. In Purple: H3K4me1 peaks for the sum of the six human cell lines marks. In Green: H3K27ac peaks for the sum of the six human cell lines. In grey conservation between mammals and teleosts. In red: conserved PNREs selected for enhancer screening (in blue: non-conserved PNREs). Last rows: Individual H3K4me1 and H3K4ac marks for each human cell types: embryonic stem cells (H1hesc), mammary epithelial cells (Hmec), skeletal muscle myocites (Hsmm), umbilical vein endothelial cells (Huvec), epidermal keratinocytes (Nhek) and lung fibroblasts (Nhlf). See also STab.1 and Tab1. Adapted from UCSC Genome Browser.



Supp. Fig. 2. The *opo* mRNA expression pattern. ISH sections showing the *opo* mRNA expression pattern from 20 somites to 72hpf. Arrowheads: neural tube (nt); optic vesicle (ov); otic vesicle (otv); pronephros (p); basal retina (br); notochord (nt); cranial neural crest (cnc); ear (e).

Name	Classification			Color code
	Highly conserved	H3K4m1 H3K27ac / human cell line	H3K4m1 H3K27ac in Zebrafish 24hpf and/or 48hpf	
H6:09040	✔	*H1hesc*	✔	
H6:09360	✔	-	✔	
H6:09510	✔	*Hmec;Nhek; Nhlf;*	✔	
H6:09516	✔	*H1hesc; Hmec; Nhlf*		
H6:09536	✔	*Hmec*		
H6:09563	✔	*H1hesc*	✔✔	
H6:09601	✔	-	✔✔	
H6:09617	✔	*Hsmm*	✔✔	
H6:09628		*H1hesc; Hmec; Hsmm; Huvec; Nhlf ;Hmec; Hsmm*	-	
H6: 09647	✔	*Hmec; Huvec; Nhlf;*		
H6:09671	✔	*Huvec*		
H6:09742	✔	-	✔✔	
H6:09778	✔	-		
H6:09813	✔	-	✔	
H6:09815	✔	-		
H6:09892	✔	*H1hesc*		
H6:09941		*Huvec; Huvec*		
H6:09949	✔	*H1hesc*	✔✔	
H6:10043	✔	*Hmec; Nhek*	✔✔	
H6:10126	✔	*H1hesc*	✔✔	
H6:10137	✔	*Hmec*	✔✔	
H6:10258	✔	*Hmec; Nhek*		
H6:10309		*H1hesc; Hmec; Nhek; Hmec; Nhek*	-	
H6:10377	✔	*H1hesc; Hmec; Nhek*	✔✔	
H6:10422	✔	*H1hesc; Hmec; Nhek; Nhlf; Hmec*	-	
H6:10431	✔	*H1hesc; Hmec; Nhek*	✔✔	
H6:10455		*Hmec; Nhek; Hmec; Nhek*	-	

Supp. Tab.1. PNREs characterization. Last column color code- Orange: conserved elements with H3K4m1 marks, in human or zebrafish. Red: conserved elements with H3K4m1 and H3K27ac marks, in human or zebrafish. Brown: non-conserved elements with H3K4m1 and H3K27ac marks in human cell lines. Yellow: conserved elements with any epigenetic mark for enhancer activity.

Name	Localization in Human genome(hg18)			Localization in Zebrafish genome (Zv8)			24hpf		48hpf	
	chr	start	end	chr	start	end	H3K4m1	H3K27ac	H3K4m1	H3K27ac
H6:09040	chr6	9040802	9041495	chr24	8345698	8345778	✔			
H6:09360	chr6	9360754	9362917	Zv8_NA8792	110063	110283	✔		✔	
H6:09510	chr6	9510079	9512442	Zv8_NA543	17616	21052	✔			
H6:09516	chr6	9516753	9517809	Zv8_NA8792	40759	40959				
H6:09536	chr6	9536004	9537930	Zv8_NA8792	31672	31862				
H6:09563	chr6	9563974	9565999	Zv8_NA8792	28994	29489	✔	✔	✔	✔
H6:09601	chr6	9601172	9602202	Zv8_NA8792	14180	14721	✔	✔	✔	✔
H6:09617	chr6	9617863	9618922	Zv8_NA8792	10854	11246			✔	✔
H6: 09647	chr6	9647690	9648627	Zv8_NA2790	5557	5998				
H6:09671	chr6	9671795	9672798	Zv8_NA2790	14478	14714	✔	✔		
H6:09742	chr6	9742877	9745652	Zv8_NA3223	14038	14237			✔	✔
H6:09778	chr6	9778300	9780223	Zv8_NA3223	29292	30058				
H6:09813	chr6	9813174	9814241	Zv8_NA3223	47089	47755			✔	
H6:09815	chr6	9815794	9816729	Zv8_NA1820	29547	29752				
H6:09892	chr6	9892240	9894126	Zv8_NA3223	51807	86254				
H6:09949	chr6	9949629	9950467	Zv8_NA3223	116792	117372	✔		✔	✔
H6:10043	chr6	10043446	10045884	Zv8_NA3223	145737	146467	✔	✔	✔	✔
H6:10126	chr6	10126474	10127046	Zv8_NA941	474	990	✔	✔		
H6:10137	chr6	10137953	10139056	Zv8_NA941	4522	5455	✔	✔	✔	✔
H6:10258	chr6	10258016	10259047	chr24	8404701	8405498				
H6:10377	chr6	10377252	10378117	chr24	8433358	8433511	✔	✔		
H6:10422	chr6	10422720	10424962	chr24	849730	8497649				
H6:10431	chr6	10431893	10433099	chr24	8470504	8470825	✔	✔	✔	✔

Supp. Tab. 2. Localization of the conserved PNREs in human and zebrafish genomes. Annotation of epigenetic marks for enhancer activity based on the data available in (Bogdanovic et al., 2012a).

PNREs Name	n.º Founders analysed	n.º Founders with RFP transmission (transgenisis control)	n.º Founders with GFP transmission (PNRE activity)	GFP activity
H6:09040	**13**	**3**	**3**	**Yes**
H6:09360	19	3	0	No
H6:09510	20	4	0	No
H6:09516	**23**	**5**	**5**	**Yes**
H6:09536	10	3	0	No
H6:09563	**25**	**8**	**8**	**Yes**
H6:09601	**13**	**3**	**3**	**Yes**
H6:09617	**13**	**3**	**3**	**Yes**
H6:09628	10	3	0	No
H6:09647	12	4	0	No
H6:09671	14	3	0	No
H6:09742	10	3	0	No
H6:09778	**16**	**7**	**5**	**Yes**
H6:09813	**20**	**4**	**2**	**Yes**
H6:09815	15	3	0	No
H6:09892	64	6	0	No
H6:09941	53	7	0	No
H6:09949	**10**	**3**	**3**	**Yes**
H6:10043	13	4	0	No
H6:10126	**20**	**7**	**4**	**Yes**
H6:10137	**12**	**6**	**6**	**Yes**
H6:10258	12	8	0	No
H6:10309	10	3	0	No
H6:10377	**41**	**12**	**12**	**Yes**
H6:10422	10	3	1	No
H6:10431	**13**	**4**	**4**	**Yes**
H6:10455	10	4	0	No

Supp. Tab. 3. PNREs transgenic assay in zebrafish. Results from the F_0 outcross and the F_1 generations. Classification of GFP activity in F_1 progeny.

Name	Strand / Position / Motif	Strand / Position / Motif	Expression Pattern
CHX10	- / 36-49 / TCTAATTA	+ / 643-656 / TAATTAGA	Neural Retina
HFH4	- / 41-53 / TAATTAAAC	+ / 639-651 / GTTTAATTA	-
NKX25	- / 77-84 / AATTATG	+ / 609-616 / CATAATT	-
NKX25L	- / 78-84 / AATTA	+ / 609-615 / TAATT	-
XBP1	- / 80-96 / TACTTCA	+ / 597-613 / TGAAGTG	Notocord; Epidermis
TITF1	- / 85-94 / ACTTCATT	+ / 599-608 / AATGAAGT	Forebrain; Neural plate
OCT1	- / 87-99 / TCATTA	+ / 594-606 / TAATGA	CNS
OCT1	+ / 92-103 / TATGTTAAAT	- / 590-601 / ATTTAGCATA	CNS
POU3F2	+ / 93-102 / TTATGTTAAA	- / 591-600 / TTTAGCATAA	CNS
CDC5	- / 93-104 / TATGTTAAATG	+ / 589-600 / CATTTAGCATA	Retina; neural crest
ZEC	+ / 100-112 / AAATGTTAGTTGC	- / 581-593 / GCAACTAACATTT	Whole organism
XFD2	+ / 128-141 / ATAAACA	- / 553-566 / TGTTTAT	-
VMYB	+ / 160-168 / AACTG	- / 526-534 / CAGTT	-
HNF1	- / 168-185 / CTCATTAAC	+ / 509-526 / GTTAATGAG	-
BACH2	- / 170-180 / TTACTCA	+ / 514-524 / TGAGTAA	-
HNF1	- / 170-184 / ACTCATTAAC	+ / 510-524 / GTTAATGAGT	Neural plate
AFP1	+ / 178-188 / ATTAACAAGAT	- / 506-516 / GTCTTGTTAAT	-
ETS2	- / 206-219 / TTCCTC	+ / 475-488 / GAGGAA	Mesoderm
ETS_Q6	+ / 210-217 / TTCC	- / 477-484 / GGAA	-
CRX	+ / 215-227 / TAA	- / 467-479 / TTA	Neural Retina
LPOLYA_B	- / 229-236 / ATTTATT	+ / 459-466 / AATAAAT	-
CDPCR1	- / 339-348 / ATCGAT	+ / 347-356 / ATCGAT	-
CDPCR3H	- / 339-348 / ATCGAT	+ / 347-356 / ATCGAT	-
GATA1	- / 402-411 / ATC	+ / 285-294 / GAT	Hematopoietic system
GFI1	+ / 427-439 / AATC	- / 257-269 / GATT	Gut; liver; pancreas
XFD3	+ / 443-456 / GTGAACA	- / 240-253 / TGTTCAC	Gut; pancreas
FOXO1	- / 443-456 / GTGAACAA	+ / 240-253 / TTGTTCAC	Gut; pancreas
FOXO1	- / 449-462 / AAAAACAA	+ / 234-247 / TTGTTTTT	Gut; pancreas
FOXO4	- / 449-462 / AAAAACAA	+ / 234-247 / TTGTTTTT	-
FOXO1	+ / 451-460 / AAACA	- / 236-245 / TGTTT	Gut; pancreas
AFP1	+ / 452-462 / AAAAACAACAC	- / 234-244 / GTGTTGTTTTT	-
FOXO4	+ / 452-462 / AAACA	- / 234-244 / TGTTT	-
BRCA	- / 457-464 / CAAC	+ / 232-239 / GTTG	CNS; eye; inner ear
MYB	+ / 460-469 / AAC	- / 227-236 / GTT	Myotome
TAACC	- / 469-491 / TTGCAGTTTGATTATTGTTT	+ / 205-227 / AAACAATAATCAAACTGCAA	-
PBX	+ / 476-487 / GTTTGA	- / 209-220 / TCAAAC	CNS; eye; swim bladder
CDP	- / 477-488 / TTTGATTATTGT	+ / 208-219 / ACAATAATCAAA	-
PBX1	- / 478-486 / TGATTATT	+ / 210-218 / AATAATCA	CNS; eye; swim bladder
SRY	- / 481-492 / TTGTT	+ / 204-215 / AACAA	-
FOXO4	- / 482-492 / TGTTT	+ / 204-214 / AAACA	-
FREAC2	- / 482-497 / TGTTTAT	+ / 199-214 / ATAAACA	-
FREAC7	- / 482-497 / TGTTTAT	+ / 199-214 / ATAAACA	-
SOX5	- / 482-491 / ATTGTT	+ / 205-214 / AACAAT	Forebrain; tail bud
XFD2	- / 482-495 / TGTTTAT	+ / 201-214 / ATAAACA	-
XFD3	- / 482-495 / TGTTTAT	+ / 201-214 / ATAAACA	-
HFH1	+ / 483-494 / ATTGTTTATAG	- / 202-213 / CTATAAACAAT	-
FOXO1	- / 484-493 / TGTTT	+ / 203-212 / AAACA	Gut; pancreas
GATA1	+ / 489-502 / GATA	- / 194-207 / TATC	Hematopietic system
NKX62	- / 534-545 / AATTA	+ / 151-162 / TAATT	-
FREAC3	- / 536-551 / ATTATTTAC	+ / 145-160 / GTAAATAAT	-
BACH2	- / 551-561 / AGACTCA	+ / 135-145 / TGAGTCT	Whole organism
OCT1	- / 559-571 / TCATTA	+ / 125-137 / TAATGA	CNS
HOXA3	+ / 578-586 / T	- / 110-118 / A	CNS; pharyngeal arch
VBP	+ / 592-601 / TTAT	- / 95-104 / ATAA	-
ZEC	+ / 599-611 / CAATGTTAGTTGA	- / 85-97 / TCAACTAACATTG	Whole organism
CDC5	+ / 610-621 / GATTTAACACA	- / 75-86 / TGTGTTAAATC	Retina; neural crest

Supp.Tab. 4. ECR Browser results for transcription factor binding sites in H6:10137 enhancer. Transcription Factors Multif alignments in the H610137 conserved sequence between humans (H6:10137 in hg19: chr6:10030070-10030764) and *xenopus tropicalis* (H6:10137 in xenTro3: GL172669:1277890-1278578). Both sequences are similar from the position 104 to 798bps (length: 689bps; sequence homology of 63%). Adapted from: ECR browser: http://ecrbrowser.dcode.org/). Last column: expression patterns at 24hpf in observed in the Zebrafish Model Organism Database (ZFIN).

Model name	Score	Start	End	Strand	Motif
Vsx2	13.491	143	151	-1	TTTAATTAG
Lim1	12.203	144	150	-1	TTAATTA
Scr	11.398	279	285	-1	TTAATGA
Scr	11.398	952	958	-1	TTAATGA
Rx	11.083	143	149	1	CTAATTA
Vsx1	10.888	144	150	-1	TTAATTA
Rx	10.866	144	150	-1	TTAATTA
otp	10.860	144	150	-1	TTAATTA
ap	10.801	143	149	1	CTAATTA
Oct	10.741	144	151	-1	TTTAATTA
exex	10.724	143	149	1	CTAATTA
en	10.719	144	150	-1	TTAATTA
exex	10.536	144	150	-1	TTAATTA
Rx	10.388	682	688	1	CTAATTG
Rx	10.388	819	825	1	CTAATTG
vsx1	10.366	143	149	1	CTAATTA
ap	10.346	144	150	-1	TTAATTA
en	10.312	143	149	1	CTAATTA
Scr	10.175	144	150	-1	TTAATTA
otp	10.093	143	149	1	CTAATTA

Supp. Tab. 5. JASPAR database transcription factor binding sites predictions in H610137 enhancer sequence. 102 putative sites were predicted with 90% sequence homology. In this table, only the results for binding sites with a socre $\geq$10 are shown.

		No eye defects (%)	Phenotipic eye defects (%)	Anoftalmia (%)	*n*
vsx2I1E2 300uM	Exp.1	17	70	13	40
	Exp.2	26	63	11	43
	Exp.3	25	64	11	45
	Mean	**22,7**	**65,7**	**11,7**	***n=128***
	St. Dev	**4,9**	**3,8**	**1,2**	
vsx2E2I2 300uM	Exp.1	22	58	1	40
	Exp.2	29	58	1	77
	Exp.3	32	54	14	87
	Mean	**27,7**	**56,7**	**5,3**	***n=204***
	St. Dev	**5,1**	**2,3**	**7,5**	
vsx2E2I2 + vsx2mRNA	Exp.1	70	29	0	53
	Exp.2	68	25	1	23
	Exp.3	53	47	0	97
	Mean	**63,7**	**33,7**	**0,3**	***n=173***
	St. Dev	**9,3**	**11,7**	**0,6**	
Controls	Exp.1	100	0	0	40
	Exp.2	100	0	0	40
	Exp.3	100	0	0	51
	Mean	**100,0**	**0,0**	**0,0**	***n=131***
	St. Dev	**0,0**	**0,0**	**0,0**	
	Exp.1	100	0	0	35
	Exp.2	100	0	0	38
	Exp.3	100	0	0	42
	Mean	**100,0**	**0,0**	**0,0**	***n=115***
	St. Dev	**0,0**	**0,0**	**0,0**	
	Exp.1	100	0	0	48
	Exp.2	100	0	0	20
	Exp.3	100	0	0	79
	Mean	**100,0**	**0,0**	**0,0**	***n=147***
	St. Dev	**0,0**	**0,0**	**0,0**	

Supp. Tab. 6. Percentage of phenotypic malformations and anoftalmia in *vsx2* morphant embryos. Data obtained from several *vsx2*MO injections, using the Tg(*vsx2.2*GFP:caax) line. St. Dev: Standard Deviation of the Mean.

		Microftalmia - Normalized eye length at 24hpf (% embryos)			
		4 to <6	6 to <8	8 to <10	*n*
vsx2I1E2mo 300uM	Exp1	45	40	15	40
	Exp2	15	76	9	47
	Exp3	19	71	10	21
	Mean	26,3	62,3	11,3	*n*=108
	St. Dev	16,3	19,5	3,2	
vsx2E2I2mo 300uM	Exp1	41	56	3	32
	Exp2	29	59	12	17
	Exp3	28	58	15	40
	Mean	32,7	57,7	10,0	*n*=89
	St. Dev	7,2	1,5	6,2	
Controls	Exp1	0	37	63	40
	Exp2	0	21	79	34
	Exp3	0	0	100	13
	Mean	0	19,3	80,7	*n*=87
	St. Dev	0	18,6	18,6	
	Exp1	0	32	68	31
	Exp2	0	38	62	13
	Exp3	0	26	74	40
	Mean	0	32,0	68,0	*n*=84
	St. Dev	0	6,0	6,0	

Supp. Tab. 7. Microftalmia percentages in *vsx2* morphant embryos. Data obtained from several *vsx2*MOs injections in Tg(*vsx2.2*GFP:caax) line. St. Dev: Standard Deviation of the Mean.

		Ventral opening angle at 24hpf (% embryos)					
		0º - 20º	21º - 40º	41º - 60º	61º - 80º	>80º	*n*
vsx2I1E2 300uM	Exp1	29	0	12	47	12	17
	Exp2	16	24,4	17,6	31	11	45
	Exp3	20,8	6,1	21	29,1	23	48
	Mean	**21,9**	**10,2**	**16,9**	**35,7**	**15,3**	***n*=110**
	St. Dev	**6,6**	**12,7**	**4,5**	**9,8**	**6,7**	
vsx2E2I2 300uM	Exp1	27	32	5	13	23	40
	Exp2	29	12,9	16	22,1	20	31
	Exp3	41	19	10	10	20	58
	Mean	**32,3**	**21,3**	**10,3**	**15,0**	**21,0**	***n*=129**
	St. Dev	**6,2**	**8,0**	**4,5**	**5,1**	**1,4**	
Controls	Exp1	100	0	0	0	0	17
	Exp2	100	0	0	0	0	45
	Exp3	100	0	0	0	0	31
	Mean	**100**	**0**	**0**	**0**	**0**	***n*=93**
	St. Dev	**0**	**0**	**0**	**0**	**0**	
	Exp1	100	0	0	0	0	40
	Exp2	100	0	0	0	0	31
	Exp3	100	0	0	0	0	42
	Mean	**100**	**0**	**0**	**0**	**0**	***n*=113**
	St. Dev	**0**	**0**	**0**	**0**	**0**	

Supp. Tab. 8. Percentage of ventral opening angles in *vsx2* morphant embryos. Data obtained from multiple *vsx2*MOs injections in Tg(*vsx2.2*GFP:caax) line. St. Dev: Standard Deviation of the Mean.

IIIV. References

Adelmann, H. B. (1929) 'Experimental studies on the development of the eye. I. The effect of the removal of median and lateral areas of the anterior end of the urodelan neural plate on the development of the eyes (*Triton teniatus* and *Amblystoma punctatum*)', *Journal of Experimental Zoology* 54: 249-290.

Adler, R. and Belecky-Adams, T. L. (2002) 'The role of bone morphogenetic proteins in the differentiation of the ventral optic cup', *Development* 129(13): 3161-71.

Adler, R. and Canto-Soler, M. V. (2007) 'Molecular mechanisms of optic vesicle development: complexities, ambiguities and controversies', *Dev Biol* 305(1): 1-13.

Banerji, J., Rusconi, S. and Schaffner, W. (1981) 'Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences', *Cell* 27(2 Pt 1): 299-308.

Bannister, A. J. and Kouzarides, T. (2011) 'Regulation of chromatin by histone modifications', *Cell Res* 21(3): 381-95.

Bar-Yosef, U., Abuelaish, I., Harel, T., Hendler, N., Ofir, R. and Birk, O. S. (2004) 'CHX10 mutations cause non-syndromic microphthalmia/ anophthalmia in Arab and Jewish kindreds', *Hum Genet* 115(4): 302-9.

Barabino, S. M. L., Spada, F., Cotelli, F. and Boncinelli, E. (1997) 'Inactivation Of the Zebrafish Homologue Of Chx10 By Antisense Oligonucleotides Causes Eye Malformations Similar to the Ocular Retardation Phenotype', *Mechanisms of Development* 63(2): 133-143.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W. J., Mattick, J. S. and Haussler, D. (2004) 'Ultraconserved elements in the human genome', *Science* 304(5675): 1321-5.

Belecky-Adams, T., Tomarev, S., Li, H. S., Ploder, L., McInnes, R. R., Sundin, O. and Adler, R. (1997) 'Pax-6, Prox 1, and Chx10 homeobox gene expression correlates with phenotypic fate of retinal precursor cells', *Investigative Ophthalmology & Visual Science* 38(7): 1293-1303.

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K. et al. (2006) 'A bivalent chromatin structure marks key developmental genes in embryonic stem cells', *Cell* 125(2): 315-26.

Bessa, J. and Gomez-Skarmeta, J. L. (2011) 'Making reporter gene constructs to analyze cis-regulatory elements', *Methods Mol Biol* 772: 397-408.

Bessa, J., Tena, J. J., de la Calle-Mustienes, E., Fernandez-Minan, A., Naranjo, S., Fernandez, A., Montoliu, L., Akalin, A., Lenhard, B., Casares, F. et al. (2009) 'Zebrafish enhancer detection (ZED) vector: a new tool to facilitate transgenesis and the functional analysis of cis-regulatory regions in zebrafish', *Dev Dyn* 238(9): 2409-17.

Bharti, K., Nguyen, M. T., Skuntz, S., Bertuzzi, S. and Arnheiter, H. (2006) 'The other pigment cell: specification and development of the pigmented epithelium of the vertebrate eye', *Pigment Cell Res* 19(5): 380-94.

Bogdanovic, O., Delfino-Machin, M., Nicolas-Perez, M., Gavilan, M. P., Gago-Rodrigues, I., Fernandez-Minan, A., Lillo, C., Rios, R. M., Wittbrodt, J. and Martinez-Morales, J. R.

(2012a) 'Numb/Numbl-Opo antagonism controls retinal epithelium morphogenesis by regulating integrin endocytosis', *Dev Cell* 23(4): 782-95.

Bogdanovic, O., Fernandez-Minan, A., Tena, J. J., de la Calle-Mustienes, E., Hidalgo, C., van Kruysbergen, I., van Heeringen, S. J., Veenstra, G. J. and Gomez-Skarmeta, J. L. (2012b) 'Dynamics of enhancer chromatin signatures mark the transition from pluripotency to cell specification during embryogenesis', *Genome Res* 22(10): 2043-53.

Bryant, D. M. and Mostov, K. E. (2008) 'From cells to organs: building polarized tissue', *Nat Rev Mol Cell Biol* 9(11): 887-901.

Bryne, J. C., Valen, E., Tang, M. H., Marstrand, T., Winther, O., da Piedade, I., Krogh, A., Lenhard, B. and Sandelin, A. (2008) 'JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update', *Nucleic Acids Res* 36(Database issue): D102-6.

Burmeister, M., Novak, J., Liang, M.-Y., Basu, S., Ploder, L., Hawes, N. L., Vidgen, D., Hoover, F., Goldman, D., Kalnins, V. I. et al. (1996) 'Ocular retardation mouse caused by *Chx10* homeobox null allele: impaired retinal progenitor proliferation and bipolar cell differentiation', *Nature Genetics* 12: 376-384.

Burridge, K. and Chrzanowska-Wodnicka, M. (1996) 'Focal adhesions, contractility, and signaling', *Annu Rev Cell Dev Biol* 12: 463-518.

Cai, Z., Feng, G. S. and Zhang, X. (2010) 'Temporal requirement of the protein tyrosine phosphatase Shp2 in establishing the neuronal fate in early retinal development', *J Neurosci* 30(11): 4110-9.

Calderwood, D. A., Fujioka, Y., de Pereda, J. M., Garcia-Alvarez, B., Nakamoto, T., Margolis, B., McGlade, C. J., Liddington, R. C. and Ginsberg, M. H. (2003) 'Integrin beta cytoplasmic domain interactions with phosphotyrosine-binding domains: a structural prototype for diversity in integrin signaling', *Proc Natl Acad Sci U S A* 100(5): 2272-7.

Cavodeassi, F., Carreira-Barbosa, F., Young, R. M., Concha, M. L., Allende, M. L., Houart, C., Tada, M. and Wilson, S. W. (2005) 'Early stages of zebrafish eye formation require the coordinated activity of Wnt11, Fz5, and the Wnt/beta-catenin pathway', *Neuron* 47(1): 43-56.

Chauhan, B. K., Disanza, A., Choi, S. Y., Faber, S. C., Lou, M., Beggs, H. E., Scita, G., Zheng, Y. and Lang, R. A. (2009) 'Cdc42- and IRSp53-dependent contractile filopodia tether presumptive lens and retina to coordinate epithelial invagination', *Development* 136(21): 3657-67.

Chen, C. M. and Cepko, C. L. (2000) 'Expression of Chx10 and Chx10-1 in the developing chicken retina', *Mech Dev* 90(2): 293-7.

Chow, R. L. and Lang, R. A. (2001) 'Early eye development in vertebrates', *Annual Review of Cell & Developmental Biology* 17: 255-96.

Chuang, J. C. and Raymond, P. A. (2002) 'Embryonic origin of the eyes in teleost fish', *Bioessays* 24(6): 519-29.

Clark, A. M., Yun, S., Veien, E. S., Wu, Y. Y., Chow, R. L., Dorsky, R. I. and Levine, E. M. (2008) 'Negative regulation of Vsx1 by its paralog Chx10/Vsx2 is conserved in the vertebrate retina', *Brain Res* 1192: 99-113.

Comings, D. E., Comings, B. G., Devor, E. J. and Cloninger, C. R. (1984) 'Detection of major gene for Gilles de la Tourette syndrome', *Am J Hum Genet* 36(3): 586-600.

Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A. et al. (2010) 'Histone H3K27ac separates active from poised enhancers and predicts developmental state', *Proc Natl Acad Sci U S A* 107(50): 21931-6.

Davies, S. J., Wise, C., Venkatesh, B., Mirza, G., Jefferson, A., Volpi, E. V. and Ragoussis, J. (2004) 'Mapping of three translocation breakpoints associated with orofacial clefting within 6p24 and identification of new transcripts within the region', *Cytogenet Genome Res* 105(1): 47-53.

Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. (2002) 'Capturing chromosome conformation', *Science* 295(5558): 1306-11.

Dudley, A. T. and Robertson, E. J. (1997) 'Overlapping expression domains of bone morphogenetic protein family members potentially account for limited tissue defects in BMP7 deficient embryos', *Developmental Dynamics* 208: 349-362.

Eiraku, M., Takata, N., Ishibashi, H., Kawada, M., Sakakura, E., Okuda, S., Sekiguchi, K., Adachi, T. and Sasai, Y. (2011) 'Self-organizing optic-cup morphogenesis in three-dimensional culture', *Nature* 472(7341): 51-6.

Esteve, P. and Bovolenta, P. (2006) 'Secreted inducers in vertebrate eye development: more functions for old morphogens', *Curr Opin Neurobiol* 16(1): 13-9.

Ezratty, E. J., Bertaux, C., Marcantonio, E. E. and Gundersen, G. G. (2009) 'Clathrin mediates integrin endocytosis for focal adhesion disassembly in migrating cells', *J Cell Biol* 187(5): 733-47.

Feijen, A., Goumans, M. J. and van den Eijnden-van Raaij, A. J. (1994) 'Expression of activin subunits, activin receptors and follistatin in postimplantation mouse embryos suggests specific developmental functions for different activins', *Development* 120(12): 3621-37.

Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M. and Dubchak, I. (2004) 'VISTA: computational tools for comparative genomics', *Nucleic Acids Res* 32(Web Server issue): W273-9.

Fuhrmann, S. (2010) 'Eye morphogenesis and patterning of the optic vesicle', *Curr Top Dev Biol* 93: 61-84.

Fuhrmann, S., Chow, L. and Reh, T. A. (2000) 'Molecular control of cell diversification in the vertebrate retina', *Results & Problems in Cell Differentiation* 31: 69-91.

Geng, X., Speirs, C., Lagutin, O., Inbal, A., Liu, W., Solnica-Krezel, L., Jeong, Y., Epstein, D. J. and Oliver, G. (2008) 'Haploinsufficiency of Six3 fails to activate Sonic hedgehog expression in the ventral forebrain and causes holoprosencephaly', *Dev Cell* 15(2): 236-47.

Gomez Skarmeta, J. L. (2009) 'Non-coding regulatory regions in genomes. Editorial', *Brief Funct Genomic Proteomic* 8(4): 213-4.

Gomez-Skarmeta, J. L., Lenhard, B. and Becker, T. S. (2006) 'New technologies, new findings, and new concepts in the study of vertebrate cis-regulatory sequences', *Dev Dyn* 235(4): 870-85.

Graf, T. and Enver, T. (2009) 'Forcing cells to change lineages', *Nature* 462(7273): 587-94.

Graw, J. (2010) 'Eye development', *Curr Top Dev Biol* 90: 343-86.

Green, E. S., Stubbs, J. L. and Levine, E. M. (2003) 'Genetic rescue of cell number in a mouse model of microphthalmia: interactions between Chx10 and G1-phase cell cycle regulators', *Development* 130(3): 539-52.

Hagege, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., de Laat, W. and Forne, T. (2007) 'Quantitative analysis of chromosome conformation capture assays (3C-qPCR)', *Nat Protoc* 2(7): 1722-33.

Hallmayer, J. F., Kalaydjieva, L., Badcock, J., Dragovic, M., Howell, S., Michie, P. T., Rock, D., Vile, D., Williams, R., Corder, E. H. et al. (2005) 'Genetic evidence for a distinct subtype of schizophrenia characterized by pervasive cognitive deficit', *Am J Hum Genet* 77(3): 468-76.

He, L., Wang, X., Tang, H. L. and Montell, D. J. (2010) 'Tissue elongation requires oscillating contractions of a basal actomyosin network', *Nat Cell Biol* 12(12): 1133-42.

Heintzman, N. D., Hon, G. C., Hawkins, R. D., Kheradpour, P., Stark, A., Harp, L. F., Ye, Z., Lee, L. K., Stuart, R. K., Ching, C. W. et al. (2009) 'Histone modifications at human enhancers reflect global cell-type-specific gene expression', *Nature* 459(7243): 108-12.

Heintzman, N. D., Stuart, R. K., Hon, G., Fu, Y., Ching, C. W., Hawkins, R. D., Barrera, L. O., Van Calcar, S., Qu, C., Ching, K. A. et al. (2007) 'Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome', *Nat Genet* 39(3): 311-8.

Hilfer, S. R. (1983) 'Development of the eye of the chick embryo', *Scan Electron Microsc*(Pt 3): 1353-69.

Hong, J. W., Hendrix, D. A. and Levine, M. S. (2008) 'Shadow enhancers as a source of evolutionary novelty', *Science* 321(5894): 1314.

Horsford, D. J., Nguyen, M. T., Sellar, G. C., Kothary, R., Arnheiter, H. and McInnes, R. R. (2005) 'Chx10 repression of Mitf is required for the maintenance of mammalian neuroretinal identity', *Development* 132(1): 177-87.

Hyer, J., Kuhlman, J., Afif, E. and Mikawa, T. (2003) 'Optic cup morphogenesis requires pre-lens ectoderm but not lens differentiation', *Dev Biol* 259(2): 351-63.

Hynes, R. O. (1992) 'Integrins: versatility, modulation, and signaling in cell adhesion', *Cell* 69(1): 11-25.

Jao, L. E., Wente, S. R. and Chen, W. (2013) 'Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system', *Proc Natl Acad Sci U S A* 110(34): 13904-9.

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. and Schilling, T. F. (1995) 'Stages of embryonic development of the zebrafish', *Dev Dyn* 203(3): 253-310.

Kish, P. E., Bohnsack, B. L., Gallina, D., Kasprick, D. S. and Kahana, A. (2011) 'The eye as an organizer of craniofacial development', *Genesis* 49(4): 222-30.

Kolsch, V., Seher, T., Fernandez-Ballester, G. J., Serrano, L. and Leptin, M. (2007) 'Control of Drosophila gastrulation by apical localization of adherens junctions and RhoGEF2', *Science* 315(5810): 384-6.

Kwan, K. M., Fujimoto, E., Grabher, C., Mangum, B. D., Hardy, M. E., Campbell, D. S., Parant, J. M., Yost, H. J., Kanki, J. P. and Chien, C. B. (2007) 'The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs', *Dev Dyn* 236(11): 3088-99.

Kwan, K. M., Otsuna, H., Kidokoro, H., Carney, K. R., Saijoh, Y. and Chien, C. B. (2012) 'A complex choreography of cell movements shapes the vertebrate eye', *Development* 139(2): 359-72.

Langheinrich, U., Hennen, E., Stott, G. and Vacun, G. (2002) 'Zebrafish as a model organism for the identification and characterization of drugs and genes affecting p53 signaling', *Curr Biol* 12(23): 2023-8.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R. et al. (2007) 'Clustal W and Clustal X version 2.0', *Bioinformatics* 23(21): 2947-8.

Letizia, A., Sotillos, S., Campuzano, S. and Llimargas, M. (2011) 'Regulated Crb accumulation controls apical constriction and invagination in Drosophila tracheal cells', *J Cell Sci* 124(Pt 2): 240-51.

Li, H., Tierney, C., Wen, L., Wu, J. Y. and Rao, Y. (1997) 'A single morphogenetic field gives rise to two retina primordia under the inluence of the prechordal plate', *Development* 124: 603-615.

Li, Z., Joseph, N. M. and Easter, S. S. J. (2000) 'The morphogenesis of the zebrafish eye, including a fate map of the optic vesicle', *Developmental Dynamics* 218: 175-188.

Liu, I. S. C., Chen, J. D., Ploder, L., Vidgen, D., Vanderkooy, D., Kalnins, V. I. and Mcinnes, R. R. (1994) 'Developmental expression of a novel murine homeobox gene (Chx10): Evidence for roles in determination of the neuroretina and inner nuclear layer', *Neuron* 13(2): 377-393.

Loosli, F., Del Bene, F., Quiring, R., Rembold, M., Martinez-Morales, J. R., Carl, M., Grabher, C., Iquel, C., Krone, A., Wittbrodt, B. et al. (2004) 'Mutations affecting retina development in Medaka', *Mech Dev* 121(7-8): 703-14.

Loosli, F., Staub, W., Finger-Baier, K., Ober, E., Verkade, H., Wittbrodt, J. and Baier, H. (2003) 'Loss of eyes in zebrafish caused by mutation of chokh/rx3', *EMBO Reports* 4: 894-899.

Martinez-Morales, J. R., Henrich, T., Ramialison, M. and Wittbrodt, J. (2007) 'New genes in the evolution of the neural crest differentiation program', *Genome Biol* 8(3): R36.

Martinez-Morales, J. R., Rembold, M., Greger, K., Simpson, J. C., Brown, K. E., Quiring, R., Pepperkok, R., Martin-Bermudo, M. D., Himmelbauer, H. and Wittbrodt, J. (2009) 'ojoplano-mediated basal constriction is essential for optic cup morphogenesis', *Development* 136(13): 2165-75.

Martinez-Morales, J. R., Rodrigo, I. and Bovolenta, P. (2004) 'Eye development: a view from the retina pigmented epithelium', *Bioessays* 26(7): 766-77.

Martinez-Morales, J. R., Signore, M., Acampora, D., Simeone, A. and Bovolenta, P. (2001) 'Otx genes are required for tissue specification in the developing eye', *Development* 128(11): 2019-30.

Martinez-Morales, J. R. and Wittbrodt, J. (2009) 'Shaping the vertebrate eye', *Curr Opin Genet Dev* 19(5): 511-7.

Mertes, F., Martinez-Morales, J. R., Nolden, T., Sporle, R., Wittbrodt, J., Lehrach, H. and Himmelbauer, H. (2009) 'Cloning of mouse ojoplano, a reticular cytoplasmic protein expressed during embryonic development', *Gene Expr Patterns* 9(8): 562-7.

Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T. K., Koche, R. P. et al. (2007) 'Genome-wide maps of chromatin state in pluripotent and lineage-committed cells', *Nature* 448(7153): 553-60.

Murray, J. C. and Schutte, B. C. (2004) 'Cleft palate: players, pathways, and pursuits', *J Clin Invest* 113(12): 1676-8.

Neto, A., Mercader, N. and Gomez-Skarmeta, J. L. (2012) 'The Osr1 and Osr2 genes act in the pronephric anlage downstream of retinoic acid signaling and upstream of Wnt2b to maintain pectoral fin development', *Development* 139(2): 301-11.

Nguyen, M. T. and Arnheiter, H. (2000) 'Signaling and transcriptional regulation in early mammalian eye development: a link between FGF and MITF', *Development* 127: 3581-3591.

Nishimura, T. and Kaibuchi, K. (2007) 'Numb controls integrin endocytosis for directional cell migration with aPKC and PAR-3', *Dev Cell* 13(1): 15-28.

Noordermeer, D., Leleu, M., Splinter, E., Rougemont, J., De Laat, W. and Duboule, D. (2011) 'The dynamic architecture of Hox gene clusters', *Science* 334(6053): 222-5.

Osipov, V. V. and Vakhrusheva, M. P. (1983) '[Variation in the expressivity of the ocular retardation gene in mice]', *Tsitol Genet* 17(4): 39-43.

Ovcharenko, I., Nobrega, M. A., Loots, G. G. and Stubbs, L. (2004) 'ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes', *Nucleic Acids Res* 32(Web Server issue): W280-6.

Picker, A., Brennan, C., Reifers, F., Böhli, H., Holder, N. and Brand, M. (1999) 'Requirement for zebrafish *acerebellar*/FGF8 in midbrain polarization, mapping and confinement of the retinotectal projection', *Development* 126: 2967-2978.

Rembold, M., Loosli, F., Adams, R. J. and Wittbrodt, J. (2006) 'Individual cell migration serves as the driving force for optic vesicle evagination', *Science* 313(5790): 1130-4.

Rowan, S., Chen, C. M., Young, T. L., Fisher, D. E. and Cepko, C. L. (2004) 'Transdifferentiation of the retina into pigmented cells in ocular retardation mice defines a new function of the homeodomain gene Chx10', *Development* 131(20): 5139-52.

Rozen, S. and Skaletsky, H. (2000) 'Primer3 on the WWW for general users and for biologist programmers', *Methods Mol Biol* 132: 365-86.

Sandelin, A., Wasserman, W. W. and Lenhard, B. (2004) 'ConSite: web-based prediction of regulatory elements using cross-species comparison', *Nucleic Acids Res* 32(Web Server issue): W249-52.

Santagati, F. and Rijli, F. M. (2003) 'Cranial neural crest and the building of the vertebrate head', *Nat Rev Neurosci* 4(10): 806-18.

Schier, A. F., Neuhauss, S. C., Helde, K. A., Talbot, W. S. and Driever, W. (1997) 'The one-eyed pinhead gene functions in mesoderm and endoderm formation in zebrafish and interacts with no tail', *Development* 124(2): 327-42.

Schwarz, M., Cecconi, F., Bernier, G., Andrejewski, N., Kammandel, B., Wagner, M. and Gruss, P. (2000) 'Spatial specification of mammalian eye territories by reciprocal transcriptional repression of Pax2 and Pax6', *Development* 127: 4325-4334.

Shimeld, S. M. and Holland, P. W. (2000) 'Vertebrate innovations', *Proc Natl Acad Sci U S A* 97(9): 4449-52.

Smith, A. N., Miller, L. A., Radice, G., Ashery-Padan, R. and Lang, R. A. (2009) 'Stage-dependent modes of Pax6-Sox2 epistasis regulate lens development and eye morphogenesis', *Development* 136(17): 2977-85.

Splinter, E., de Wit, E., van de Werken, H. J., Klous, P. and de Laat, W. (2012) 'Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation', *Methods* 58(3): 221-30.

Sprague, J., Clements, D., Conlin, T., Edwards, P., Frazer, K., Schaper, K., Segerdell, E., Song, P., Sprunger, B. and Westerfield, M. (2003) 'The Zebrafish Information Network (ZFIN): the zebrafish model organism database', *Nucleic Acids Res* 31(1): 241-3.

Stigloher, C., Ninkovic, J., Laplante, M., Geling, A., Tannhauser, B., Topp, S., Kikuta, H., Becker, T. S., Houart, C. and Bally-Cuif, L. (2006) 'Segregation of telencephalic and eye-field identities inside the zebrafish forebrain territory is controlled by Rx3', *Development* 133(15): 2925-35.

Stormo, G. D. (2000) 'DNA binding sites: representation and discovery', *Bioinformatics* 16(1): 16-23.

Straub, R. E., MacLean, C. J., Ma, Y., Webb, B. T., Myakishev, M. V., Harris-Kerr, C., Wormley, B., Sadek, H., Kadambi, B., O'Neill, F. A. et al. (2002) 'Genome-wide scans of three independent sets of 90 Irish multiplex schizophrenia families and follow-up of selected regions in all families provides evidence for multiple susceptibility genes', *Mol Psychiatry* 7(6): 542-59.

Sundaram, S. K., Huq, A. M., Sun, Z., Yu, W., Bennett, L., Wilson, B. J., Behen, M. E. and Chugani, H. T. (2011) 'Exome sequencing of a pedigree with Tourette syndrome or chronic tic disorder', *Ann Neurol* 69(5): 901-4.

Truslove, G. M. (1962) 'A gene causing ocular retardation in the mouse', *J Embryol Exp Morphol* 10: 652-60.

Varga, Z. M., Wegner, J. and Westerfield, M. (1999) 'Anterior movement of ventral diencephalic precursors separates the primordial eye field in the neural plate and requires cyclops', *Development* 126(24): 5533-46.

Vitorino, M., Jusuf, P. R., Maurus, D., Kimura, Y., Higashijima, S. and Harris, W. A. (2009) 'Vsx2 in the zebrafish retina: restricted lineages through derepression', *Neural Dev* 4: 14.

Vogel-Hopker, A., Momose, T., Rohrer, H., Yasuda, K., Ishihara, L. and Rapaport, D. H. (2000) 'Multiple functions of fibroblast growth factor-8 (FGF-8) in chick eye development', *Mech Dev* 94(1-2): 25-36.

Wasserman, W. W. and Sandelin, A. (2004) 'Applied bioinformatics for the identification of regulatory elements', *Nat Rev Genet* 5(4): 276-87.

Wilson, S. W. and Houart, C. (2004) 'Early steps in the development of the forebrain', *Dev Cell* 6(2): 167-81.

Winkler, S., Loosli, F., Henrich, T., Wakamatsu, Y. and Wittbrodt, J. (2000) 'The conditional medaka mutation eyeless uncouples patterning and morphogenesis of the eye', *Development* 127(9): 1911-1919.

Woolfe, A., Goodson, M., Goode, D. K., Snell, P., McEwen, G. K., Vavouri, T., Smith, S. F., North, P., Callaway, H., Kelly, K. et al. (2005) 'Highly conserved non-coding sequences are associated with vertebrate development', *PLoS Biol* 3(1): e7.

Xiao, A., Wang, Z., Hu, Y., Wu, Y., Luo, Z., Yang, Z., Zu, Y., Li, W., Huang, P., Tong, X. et al. (2013) 'Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish', *Nucleic Acids Res* 41(14): e141.

Yun, S., Saijoh, Y., Hirokawa, K. E., Kopinke, D., Murtaugh, L. C., Monuki, E. S. and Levine, E. M. (2009) 'Lhx2 links the intrinsic and extrinsic factors that control optic cup formation', *Development* 136(23): 3895-906.

Zou, C. and Levine, E. M. (2012) 'Vsx2 controls eye organogenesis and retinal progenitor identity via homeodomain and non-homeodomain residues required for high affinity DNA binding', *PLoS Genet* 8(9): e1002924.

Acknowledgements

Os primeiros nesta lista de agradecimentos são e sempre o serão a minha família. Mãe, Pai e Ana Lúcia, obrigado por todo o apoio e carinho ao longo destes 4 anos (e de toda a minha vida), sem vocês isto não teria sido possível. Obrigado por apoiarem as minhas ideias loucas, que neste caso em especial, até acabou bem!

Obrigado ao meu querido Bruno Silva, o meu companheiro em tudo. Nesta luta tu desempenhaste um papel especial e tu sabes disso! Obrigado por me aturares nos bons e maus momentos, por me obrigares a descansar, por me ajudares com tudo, em casa no lab, onde quer que fosse.

Gracias a todos en el grupo de JL Skarmeta y en el grupo Juan Martinez que me ayudaran a aprender, a investigar, a leer , a ser critica, a ser autónoma y a seguir para adelante. Vosotros me formaran como doctora y vos agradezco toda la paciencia, cariño y apoyo a lo largo de estos 4 años. En especial á: Ana Neto, JL Royo, Renata, Zé, Ozren, Ana Miñan, Juan Tena, Mariana (Mariana a ti te debo gracias enorme por todo el tiempo que has dedicado á la construcción de esta tesis! Gracias por enseñar me a escribir a decent english!).

Maria, Rocio y Ozren, gracias especialmente por aguantarme diariamente (sabéis de lo que hablo!). Gracias por vuestra amistad, confianza y cariño. Maria Angeles, Adela, Manolo, Mario, Calitó y Ariza lo mismo, vos agradezco vuestra amistad y cariño. Todos vosotros habéis contribuido determinadamente para que lo haya pasado bien aquí.

Gracias a los jefes JL Skarmeta y en especial a Juan Ramón por la oportunidad que me habéis dado de me formar doctora, por todo lo que me enseñasteis, por toda la confianza y motivación que siempre supisteis transmitirme a lo largo de este tiempo.

The End